

58447

From: Chan, Christina
Sent: Wednesday, January 16, 2002 1:54 PM
To: Li, Janice; STIC-Biotech/ChemLib
Subject: RE: rush search for 09/242,202

Please rush. Thanks Chris

-----Original Message-----

Fr m: Li, Janice
Sent: Wednesday, January 16, 2002 1:26 PM
T : Chan, Christina
Subject: rush search for 09/242,202

Hi, Chris;

Could you please give approval for rush search SEQ ID Nos: 10, 16, 22, 27, 28; residues 1-12 of SEQ ID No: 22; and residues 1-11 of SEQ ID No: 10 against all commercial databases including pending patent database.

This is an amended case.

Thanks,

Q. Janice Li
Patent Examiner
AU 1632
CM1, Rm12D11
Mail Box12E12
703-308-7942
janice.li@uspto.gov

Point of Contact:
Mona Smith
Technical Info. Specialist
CM1 12C14 Tel: 308-3278

RECEIVED
JAN 16 2002
STIC

Searcher: M. Smith
Phone: _____
Location: _____
Date Picked Up: 1/17/02
Date Completed: 1/18
Searcher Prep/Review: 15
Clerical: _____
Online time: 15

TYPE OF SEARCH:
NA Sequences: 7
AA Sequences: _____
Structures: _____
Bibliographic: _____
Litigation: _____
Full text: _____
Patent Family: _____
Other: _____

VENDOR/COST(where applic.)
STN: _____
DIALOG: _____
Questel/Orbit: _____
DRLink: _____
Lexis/Nexis: _____
Sequence Sys.: _____
WWW/Internet: _____
Other (specify): _____

THIS PAGE BLANK (USPTO)

10/11/1994
10/11/1994
10/11/1994
10/11/1994

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: January 17, 2002, 11:54:39 ; Search time 221.34 Seconds
(without alignments)
12.279 Million cell updates/sec

Title: US-09-242-202a-22_COPY_1_12
Perfect score: 12
Sequence: 1 GCCACCATGGCC.12

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 351203 seqs, 113238999 residues

Total number of hits satisfying chosen parameters: 702406

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued Patents.NA.*
1: /cgn2.6/ptodata/2/ina/5A_COMB.seq.*
2: /cgn2.6/ptodata/2/ina/5B_COMB.seq.*
3: /cgn2.6/ptodata/2/ina/6A_COMB.seq.*
4: /cgn2.6/ptodata/2/ina/6B_COMB.seq.*
5: /cgn2.6/ptodata/2/ina/PCTUS_COMB.seq.*
6: /cgn2.6/ptodata/2/ina/backfiles1.seq.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	12	100.0	15	4	US-09-309-382-9
2	12	100.0	18	4	US-09-309-382-8
3	12	100.0	22	4	US-09-309-382-24
4	12	100.0	22	4	US-09-309-382-25
5	12	100.0	27	1	US-08-013-801-8
6	12	100.0	27	1	US-08-072-063-17
7	12	100.0	27	1	US-08-212-132-8
8	12	100.0	27	1	US-08-064-693-17
9	12	100.0	27	1	US-08-430-417-8
10	12	100.0	27	1	US-08-470-366-8
11	12	100.0	27	1	US-08-466-822-8
12	12	100.0	27	2	US-08-704-504-8
13	12	100.0	27	4	US-08-885-366-17
14	12	100.0	27	4	US-09-223-342-8
15	12	100.0	27	5	PCT-US93-04754-17
16	12	100.0	27	5	PCT-US94-01235-8
17	12	100.0	27	5	PCT-US95-03125-8
18	12	100.0	28	1	US-08-227-536-13
19	12	100.0	28	4	US-09-309-382-17
20	12	100.0	28	4	US-09-309-382-18
21	12	100.0	28	5	PCT-US95-04682-13
22	12	100.0	32	3	US-08-987-418A-9
23	12	100.0	32	4	US-09-343-062-9
24	12	100.0	33	2	US-08-478-386A-4
25	12	100.0	33	2	US-08-292-597-4
26	12	100.0	33	2	US-08-388-653-4
27	12	100.0	33	2	US-08-473-985-4

28	12	100.0	33	2	US-08-483-898-4	Sequence 4, Appli
29	12	100.0	33	3	US-09-087-716-4	Sequence 4, Appli
30	12	100.0	33	3	US-09-157-753-4	Sequence 4, Appli
31	12	100.0	33	3	US-09-157-230-4	Sequence 4, Appli
32	12	100.0	33	3	US-09-087-811-4	Sequence 4, Appli
33	12	100.0	33	3	US-09-156-855-4	Sequence 4, Appli
34	12	100.0	33	3	US-09-158-010-4	Sequence 4, Appli
35	12	100.0	33	4	US-09-087-647-4	Sequence 4, Appli
36	12	100.0	34	1	US-08-264-003B-6	Sequence 6, Appli
37	12	100.0	34	3	US-08-842-234-6	Sequence 6, Appli
38	12	100.0	35	4	US-09-231-182B-3	Sequence 3, Appli
39	12	100.0	37	4	US-09-309-382-20	Sequence 20, Appli
40	12	100.0	38	4	US-09-231-182B-4	Sequence 4, Appli
41	12	100.0	39	4	US-09-309-382-21	Sequence 21, Appli
42	12	100.0	39	4	US-09-309-382-22	Sequence 22, Appli
43	12	100.0	40	2	US-08-425-684-29	Sequence 29, Appli
44	12	100.0	40	2	US-08-425-684-32	Sequence 32, Appli
45	12	100.0	40	2	US-08-675-502-29	Sequence 29, Appli

ALIGNMENTS

RESULT 1
US-09-309-382-9
; Sequence 9, Application US/09309382
; Patent No. 6291214
; GENERAL INFORMATION:
; APPLICANT: Richards, Cynthia
; APPLICANT: Weiner, Michael
; TITLE OF INVENTION: SYSTEM FOR GENERATING RECOMBINANT VIRUSES
; FILE REFERENCE: PU3481US2
; CURRENT APPLICATION NUMBER: US/09/309,382
; CURRENT FILING DATE: 1999-05-10
; EARLIER APPLICATION NUMBER: 60/084,936
; EARLIER FILING DATE: 1998-05-11
; NUMBER OF SEQ ID NOS: 26
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 9
; LENGTH: 15
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: n bases may be A, T, C, G, unknown or other
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: synthetic
; OTHER INFORMATION: oligonucleotide
US-09-309-382-9

Query Match 100.0%; Score 12; DB 4; Length 15;
Best Local Similarity 100.0%; Pred. No. 6.7e+02;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GCCACCATGGCC 12
| | | | | | | | | | | | | | | | | | | | | |
Db 4 gccaccatggcc 15

RESULT 2
US-09-309-382-8
; Sequence 8, Application US/09309382
; Patent No. 6291214
; GENERAL INFORMATION:
; APPLICANT: Richards, Cynthia
; APPLICANT: Weiner, Michael
; TITLE OF INVENTION: SYSTEM FOR GENERATING RECOMBINANT VIRUSES
; FILE REFERENCE: PU3481US2
; CURRENT APPLICATION NUMBER: US/09/309,382
; CURRENT FILING DATE: 1999-05-10
; EARLIER APPLICATION NUMBER: 60/084,936
; EARLIER FILING DATE: 1998-05-11
; NUMBER OF SEQ ID NOS: 26

; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 8
; LENGTH: 18
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: recognition
; OTHER INFORMATION: site
US-09-309-382-8

Query Match 100.0%; Score 12; DB 4; Length 18;
Best Local Similarity 100.0%; Pred. No. 6.7e+02;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GCCACCATGGCC 12
|||||
Db 2 gccaccatggcc 13

RESULT 3
US-09-309-382-24
; Sequence 24, Application US/09309382
; Patent No. 6291214
; GENERAL INFORMATION:
; APPLICANT: Richards, Cynthia
; APPLICANT: Weiner, Michael
; TITLE OF INVENTION: SYSTEM FOR GENERATING RECOMBINANT VIRUSES
; FILE REFERENCE: PU3481US2
; CURRENT APPLICATION NUMBER: US/09/309,382
; CURRENT FILING DATE: 1999-05-10
; EARLIER APPLICATION NUMBER: 60/084,936
; EARLIER FILING DATE: 1998-05-11
; NUMBER OF SEQ ID NOS: 26
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 24
; LENGTH: 22
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: n bases may be A, T, C, G, unknown or other
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: synthetic
; OTHER INFORMATION: oligonucleotide
US-09-309-382-24

Query Match 100.0%; Score 12; DB 4; Length 22;
Best Local Similarity 100.0%; Pred. No. 6.6e+02;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GCCACCATGGCC 12
|||||
Db 9 gccaccatggcc 20

RESULT 4
US-09-309-382-25/c
; Sequence 25, Application US/09309382
; Patent No. 6291214
; GENERAL INFORMATION:
; APPLICANT: Richards, Cynthia
; APPLICANT: Weiner, Michael
; TITLE OF INVENTION: SYSTEM FOR GENERATING RECOMBINANT VIRUSES
; FILE REFERENCE: PU3481US2
; CURRENT APPLICATION NUMBER: US/09/309,382
; CURRENT FILING DATE: 1999-05-10
; EARLIER APPLICATION NUMBER: 60/084,936
; EARLIER FILING DATE: 1998-05-11
; NUMBER OF SEQ ID NOS: 26
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 25
; LENGTH: 22

; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: n bases may be A, T, C, G, unknown or other
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: synthetic
; OTHER INFORMATION: oligonucleotide
US-09-309-382-25

Query Match 100.0%; Score 12; DB 4; Length 22;
Best Local Similarity 100.0%; Pred. No. 6.6e+02;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GCCACCATGGCC 12
|||||
Db 14 GCCACCATGGCC 3

RESULT 5
US-08-013-801-8
; Sequence 8, Application US/08013801
; Patent No. 5420019
; GENERAL INFORMATION:
; APPLICANT: Theofan, Georgia
; APPLICANT: Horwitz, Arnold
; APPLICANT: Burke, David
; APPLICANT: Baltayan, Manik
; APPLICANT: Grinna, Lynn S
; TITLE OF INVENTION: Stable Bactericidal/Permeability-
; TITLE OF INVENTION: Increasing Protein Products and Pharmaceutical
; TITLE OF INVENTION: Compositions Containing the Same
; NUMBER OF SEQUENCES: 14
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Marshall, O'Toole, Gerstein, Murray &
; ADDRESSEE: Bicknell
; STREET: Two First National Plaza
; CITY: Chicago
; STATE: Illinois
; COUNTRY: USA
; ZIP: 60603
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/013,801
; FILING DATE: 02 FEB 1993
; CLASSIFICATION: 424
; ATTORNEY/AGENT INFORMATION:
; NAME: Meyers, Thomas C.
; REGISTRATION NUMBER: P-36,989
; REFERENCE/DOCKET NUMBER: 27129/30911
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 312/346-5750
; TELEFAX: 312/346-9740
; TELEX: 25-3856
; INFORMATION FOR SEQ ID NO: 8:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 27 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: CDNA
US-08-013-801-8

Query Match 100.0%; Score 12; DB 1; Length 27;
Best Local Similarity 100.0%; Pred. No. 6.5e+02;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GCCACCATGGCC 12

SEQUENCE CHARACTERISTICS:
LENGTH: 27 base pairs
TYPE: NUCLEIC ACID
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: CDNA
US-08-064-693-17

Query Match 100.0%; Score 12; DB 1; Length 27;
Best Local Similarity 100.0%; Pred. No. 6.5e+02;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GCCACCATGGCC 12
|||||
Db 10 GCCACCATGGCC 21

RESULT 9
US-08-430-417-8
; Sequence 8, Application US/08430417
; Patent No. 5674834
; GENERAL INFORMATION:
; APPLICANT: Theofan, Georgia
; APPLICANT: Horwitz, Arnold
; APPLICANT: Burke, David
; APPLICANT: Baitalan, Manik
; APPLICANT: Grinna, Lynn S.
; TITLE OF INVENTION: Stable Bactericidal/Permeability-Increasing
; TITLE OF INVENTION: Protein Products and Pharmaceutical Compositions Containing
; NUMBER OF SEQUENCES: 14
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Marshall, O'Toole, Gerstein, Murray & Borun
; STREET: 6300 Sears Tower, 233 South Wacker Drive
; CITY: Chicago
; STATE: Illinois
; COUNTRY: USA
; ZIP: 60606-6402
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/430,417
; FILING DATE:
; CLASSIFICATION: 514
; ATTORNEY/AGENT INFORMATION:
; NAME: Meyers, Thomas C.
; REGISTRATION NUMBER: P-36,989
; REFERENCE/DOCKET NUMBER: 27129/30911
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 312/474-6300
; TELEFAX: 312/474-0448
; TELEX: 25-3856
; INFORMATION FOR SEQ ID NO: 8:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 27 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: CDNA
US-08-430-417-8

Query Match 100.0%; Score 12; DB 1; Length 27;
Best Local Similarity 100.0%; Pred. No. 6.5e+02;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GCCACCATGGCC 12
|||||
Db 10 GCCACCATGGCC 21

RESULT 10
US-08-470-366-8
; Sequence 8, Application US/08470366
; Patent No. 5703038
; GENERAL INFORMATION:
; APPLICANT: Little, Roger
; APPLICANT: Ammons, Steve
; TITLE OF INVENTION: Therapeutic Uses of Bactericidal/Permeability-
; TITLE OF INVENTION: Increasing Protein Dimer Products
; NUMBER OF SEQUENCES: 11
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Marshall, O'Toole, Gerstein, Murray & Borun
; STREET: 6300 Sears Tower, 233 South Wacker Drive
; CITY: Chicago
; STATE: Illinois
; COUNTRY: United States of America
; ZIP: 60606-6402
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/470,366
; FILING DATE:
; CLASSIFICATION: 514
; ATTORNEY/AGENT INFORMATION:
; NAME: Sharp, Jeffrey S.
; REGISTRATION NUMBER: 31,879
; REFERENCE/DOCKET NUMBER: 27129/31735
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 312/474-6300
; TELEFAX: 312/474-0448
; TELEX: 25-3856
; INFORMATION FOR SEQ ID NO: 8:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 27 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: CDNA
US-08-470-366-8

Query Match 100.0%; Score 12; DB 1; Length 27;
Best Local Similarity 100.0%; Pred. No. 6.5e+02;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GCCACCATGGCC 12
|||||
Db 10 GCCACCATGGCC 21

RESULT 11
US-08-466-822-8
; Sequence 8, Application US/08466822
; Patent No. 5827816
; GENERAL INFORMATION:
; APPLICANT: Theofan, Georgia
; APPLICANT: Horwitz, Arnold
; APPLICANT: Burke, David
; APPLICANT: Baitalan, Manik
; APPLICANT: Grinna, Lynn S.
; TITLE OF INVENTION: Stable Bactericidal/Permeability-Increasing
; TITLE OF INVENTION: Protein Products and Pharmaceutical Compositions Containing
; NUMBER OF SEQUENCES: 14
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Marshall, O'Toole, Gerstein, Murray & Borun
; STREET: 6300 Sears Tower, 233 South Wacker Drive
; CITY: Chicago

; STATE: Illinois
; COUNTRY: USA
; ZIP: 60606-6402
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/466,822
; FILING DATE:
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Meyers, Thomas C.
; REGISTRATION NUMBER: P-36,989
; REFERENCE/DOCKET NUMBER: 27129/30911
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 312/474-6300
; TELEFAX: 312/474-0448
; TELEX: 25-3856
; INFORMATION FOR SEQ ID NO: 8:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 27 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
US-08-466-822-8

Query Match 100.0%; Score 12; DB 1; Length 27;
Best Local Similarity 100.0%; Pred. No. 6.5e+02;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GCCACCATGGCC 12
|||||
DB 10 GCCACCATGGCC 21

RESULT 12
US-08-704-504-8
; Sequence 8, Application US/08704504
; Patent No. 5856302
; GENERAL INFORMATION:
; APPLICANT: Little, Roger G.
; APPLICANT: Ammons, William Steve
; TITLE OF INVENTION: Therapeutic Uses of Bactericidal/Permeability-
; TITLE OF INVENTION: Increasing Protein Dimer Products
; NUMBER OF SEQUENCES: 11
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Marshall, O'Toole, Gerstein, Murray & Borun
; STREET: 6300 Sears Tower, 233 South Wacker Drive
; CITY: Chicago
; STATE: Illinois
; COUNTRY: United States of America
; ZIP: 60606-6402
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/704,504
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/08/212,132
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Sharp, Jeffrey S.
; REGISTRATION NUMBER: 31,879
; REFERENCE/DOCKET NUMBER: 27129/31735
; TELECOMMUNICATION INFORMATION:

; TELEPHONE: 312/474-6300
; TELEFAX: 312/474-0448
; TELEX: 25-3856
; INFORMATION FOR SEQ ID NO: 8:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 27 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
US-08-704-504-8

Query Match 100.0%; Score 12; DB 2; Length 27;
Best Local Similarity 100.0%; Pred. No. 6.5e+02;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GCCACCATGGCC 12
|||||
DB 10 GCCACCATGGCC 21

RESULT 13
US-08-885-366-17
; Sequence 17, Application US/08885366
; Patent No. 6274348
; GENERAL INFORMATION:
; APPLICANT: Theofan, Georgia
; APPLICANT: Grinna, Lynn S
; APPLICANT: Horwitz, Arnold
; TITLE OF INVENTION: BPI-Immunoglobulin Fusion Proteins
; NUMBER OF SEQUENCES: 19
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Marshall, O'Toole, Gerstein, Murray &
; ADDRESSEE: Borun
; STREET: 6300 Sears Tower, 233 South Wacker
; CITY: Chicago
; STATE: Illinois
; COUNTRY: USA
; ZIP: 60606-6402
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/885,366
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/064,693
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Meyers Thomas C.
; REGISTRATION NUMBER: 36,989
; REFERENCE/DOCKET NUMBER: 30659
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 312/474-6300
; TELEFAX: 312/474-0448
; TELEX: 25-3856
; INFORMATION FOR SEQ ID NO: 17:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 27 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
US-08-885-366-17

Query Match 100.0%; Score 12; DB 4; Length 27;
Best Local Similarity 100.0%; Pred. No. 6.5e+02;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GCCACCATGGCC 12
|||||
Db 10 GCCACCATGGCC 21

RESULT 14

US-09-223-342-8
; Sequence 8, Application US/09223342
; Patent No. 6277821
; GENERAL INFORMATION:
; APPLICANT: Little, Roger G.
; APPLICANT: Ammons, William Steve
; TITLE OF INVENTION: Therapeutic Uses of Bactericidal/Permeability-
; INCREASING PROTEIN DIMER PRODUCTS
; NUMBER OF SEQUENCES: 11
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Marshall, O'Toole, Gerstein, Murray & Borun
; STREET: 6300 Sears Tower, 233 South Wacker Drive
; CITY: Chicago
; STATE: Illinois
; COUNTRY: United States of America
; ZIP: 60606-6402
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/223,342
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/704,504
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Sharp, Jeffrey S.
; REGISTRATION NUMBER: 31,879
; REFERENCE/DOCKET NUMBER: 27129/31735
; TELEPHONE: 312/474-6300
; TELEFAX: 312/474-0448
; TELEX: 25-3856
; INFORMATION FOR SEQ ID NO: 8:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 27 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
US-09-223-342-8

Query Match 100.0%; Score 12; DB 4; Length 27;
Best Local Similarity 100.0%; Pred. No. 6.5e+02;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GCCACCATGGCC 12
|||||
Db 10 GCCACCATGGCC 21

RESULT 15

PCT-US93-04754-17
; Sequence 17, Application PC/TUS9304754
; GENERAL INFORMATION:
; APPLICANT: Theofan, Georgia
; APPLICANT: Grinna, Lynn S
; APPLICANT: Horwitz, Arnold
; TITLE OF INVENTION: BPI-immunoglobulin Fusion Proteins
; NUMBER OF SEQUENCES: 19
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Marshall, O'Toole, Gerstein, Murray &

; ADDRESSEE: Borun
; STREET: 6300 Sears Tower, 233 South Wacker
; CITY: Chicago
; STATE: Illinois
; COUNTRY: USA
; ZIP: 60606-6402
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: PCT/US93/04754
; FILING DATE: 19930519
; CLASSIFICATION:
; ATTORNEY/AGENT INFORMATION:
; NAME: Meyers Thomas C.
; REGISTRATION NUMBER: 36,989
; REFERENCE/DOCKET NUMBER: 30659
; TELEPHONE: 312/474-6300
; TELEFAX: 312/474-0448
; TELEX: 25-3856
; INFORMATION FOR SEQ ID NO: 17:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 27 base pairs
; TYPE: NUCLEIC ACID
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
PCT-US93-04754-17

Query Match 100.0%; Score 12; DB 5; Length 27;
Best Local Similarity 100.0%; Pred. No. 6.5e+02;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GCCACCATGGCC 12
|||||
Db 10 GCCACCATGGCC 21

Search completed: January 17, 2002, 11:54:39
Job time: 17406 sec

THIS PAGE BLANK (USP),

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: January 17, 2002, 11:49:53 ; Search time 10436.3 Seconds
(without alignments)
18,969 Million cell updates/sec

Title: US-09-242-202a-22_COPY_1_12
Perfect score: 12
Sequence: 1 GCCACCATGCC 12

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 1472140 seqs, 8248589755 residues

Total number of hits satisfying chosen parameters: 2944280

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : GenEmbl.*

- 1: gb_ba.*
- 2: gb_htg.*
- 3: gb_in.*
- 4: gb_om.*
- 5: gb_ov.*
- 6: gb_pat.*
- 7: gb_ph.*
- 8: gb_pl.*
- 9: gb_pr.*
- 10: gb_ro.*
- 11: gb_sts.*
- 12: gb_sy.*
- 13: gb_un.*
- 14: gb_vl.*
- 15: em_ba.*
- 16: em_fun.*
- 17: em_hum.*
- 18: em_in.*
- 19: em_om.*
- 20: em_or.*
- 21: em_ov.*
- 22: em_pat.*
- 23: em_ph.*
- 24: em_pl.*
- 25: em_ro.*
- 26: em_sts.*
- 27: em_sy.*
- 28: em_un.*
- 29: em_vl.*
- 30: em_htgo_hum.*
- 31: em_htgo_inv.*
- 32: em_htgo_rod.*
- 33: em_htg_hum.*
- 34: em_htg_inv.*
- 35: em_htg_rod.*
- 36: em_htg_other.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match	Length	DB	ID	Description
1	12	100.0	22	6	AX139440	AX139440 Sequence
2	12	100.0	23	6	AX098902	AX098902 Sequence
3	12	100.0	23	6	AX099322	AX099322 Sequence
4	12	100.0	27	6	AR027668	AR027668 Sequence
5	12	100.0	27	6	AR050743	AR050743 Sequence
6	12	100.0	27	6	AX030652	AX030652 Sequence
7	12	100.0	27	6	II13573	II13573 Sequence 17
8	12	100.0	27	6	II14327	II14327 Sequence 8
9	12	100.0	27	6	IS0683	IS0683 Sequence 17
10	12	100.0	27	6	I68187	I68187 Sequence 8
11	12	100.0	27	6	I87071	I87071 Sequence 8
12	12	100.0	28	6	I62302	I62302 Sequence 13
13	12	100.0	30	6	AX179323	AX179323 Sequence
14	12	100.0	32	6	AR146549	AR146549 Sequence
15	12	100.0	33	6	AR034283	AR034283 Sequence
16	12	100.0	33	6	AR035389	AR035389 Sequence
17	12	100.0	33	6	AR050804	AR050804 Sequence
18	12	100.0	33	6	AR053809	AR053809 Sequence
19	12	100.0	33	6	AR091590	AR091590 Sequence
20	12	100.0	33	6	AR117469	AR117469 Sequence
21	12	100.0	34	6	AR141722	AR141722 Sequence
22	12	100.0	34	6	AX103386	AX103386 Sequence
23	12	100.0	34	6	I84665	I84665 Sequence 6
24	12	100.0	35	6	AX049981	AX049981 Sequence
25	12	100.0	35	6	AX077144	AX077144 Sequence
26	12	100.0	35	6	AX099890	AX099890 Sequence
27	12	100.0	35	6	AX138011	AX138011 Sequence
28	12	100.0	37	6	AX114583	AX114583 Sequence
29	12	100.0	38	6	AX088738	AX088738 Sequence
30	12	100.0	39	6	AX179597	AX179597 Sequence
31	12	100.0	40	6	AR053605	AR053605 Sequence
32	12	100.0	40	6	AR053608	AR053608 Sequence
33	12	100.0	43	6	AR131268	AR131268 Sequence
34	12	100.0	44	6	AR040911	AR040911 Sequence
35	12	100.0	44	6	AR040950	AR040950 Sequence
36	12	100.0	44	6	AR051818	AR051818 Sequence
37	12	100.0	44	6	AR051857	AR051857 Sequence
38	12	100.0	44	6	AR115954	AR115954 Sequence
39	12	100.0	44	6	AR115993	AR115993 Sequence
40	12	100.0	44	6	AR127321	AR127321 Sequence
41	12	100.0	44	6	AR127360	AR127360 Sequence
42	12	100.0	44	6	I36208	I36208 Sequence 10
43	12	100.0	48	6	A97024	A97024 Sequence 1
44	12	100.0	48	6	AX133393	AX133393 Sequence
45	12	100.0	56	6	AR040909	AR040909 Sequence

ALIGNMENTS

RESULT	1					
AX139440						
LOCUS	AX139440	22 bp	DNA		PAT	30-MAY-2001
DEFINITION	Sequence 3 from Patent EP1094109.					
ACCESSION	AX139440					
VERSION	AX139440.1	GI:14275086				
KEYWORDS	human.					
SOURCE						
ORGANISM	Homo sapiens					
	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;					
	Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.					
REFERENCE	1 (bases 1 to 22)					
AUTHORS	Harland, L.					
TITLE	Human g protein-coupled receptor, pfi-010					
JOURNAL	Patent: EP 1094109-A 3 25-APR-2001;					
FEATURES	Pfizer Limited (GB) ; PFIZER INC. (US)					
	Location/Qualifiers					
	1..22					
	/organism="Homo sapiens"					
	/db_xref="taxon:9606"					
	source					

```
BASE COUNT      6 a      9 c      5 g      2 t
ORIGIN

Query Match      100.0%; Score 12; DB 6; Length 22;
Best Local Similarity 100.0%; Pred. No. 2.4e+04;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GCCACCATGGCC 12
Db 5 GCCACCATGGCC 16

RESULT 2
AX098902
LOCUS      AX098902      23 bp      DNA
DEFINITION Sequence 27 from Patent WO0119870.
ACCESSION  AX098902
VERSION     AX098902.1 GI:13538128
KEYWORDS   .
SOURCE      synthetic construct.
ORGANISM    synthetic construct
            artificial sequence.
REFERENCE   1 (bases 1 to 23)
AUTHORS     Brown,J.P. and Bertelli,F.
TITLE       Secreted soluble _g(a)2-g(d)-2, _g(a)2-g(d)-3 or _g(a)2-g(d)-4
            calcium channel subunit polypeptides and screening assays using
            same
JOURNAL     Patent: WO 0119870-A 27 22-MAR-2001;
            WARNER-LAMBERT COMPANY (US)
FEATURES    Location/Qualifiers
            source
            1..23
            /organism="synthetic construct"
            /db_xref="taxon:32630"
            /note="primer"

BASE COUNT      2 a      10 c      9 g      2 t
ORIGIN

Query Match      100.0%; Score 12; DB 6; Length 23;
Best Local Similarity 100.0%; Pred. No. 2.4e+04;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GCCACCATGGCC 12
Db 3 GCCACCATGGCC 14

RESULT 3
AX099322
LOCUS      AX099322      23 bp      DNA
DEFINITION Sequence 27 from Patent WO0120336.
ACCESSION  AX099322
VERSION     AX099322.1 GI:13538474
KEYWORDS   .
SOURCE      synthetic construct.
ORGANISM    synthetic construct
            artificial sequence.
REFERENCE   1 (bases 1 to 23)
AUTHORS     Bertelli,F., Brown,J.P., Dissanayake,V., Suman-Chauhan,N. and
            Gee,N.S.
TITLE       Screening for alpha2delta-1 subunit binding ligands
JOURNAL     Patent: WO 0120336-A 27 22-MAR-2001;
            WARNER-LAMBERT COMPANY (US)
FEATURES    Location/Qualifiers
            source
            1..23
            /organism="synthetic construct"
            /db_xref="taxon:32630"
            /note="primer"

BASE COUNT      2 a      10 c      9 g      2 t
ORIGIN

Query Match      100.0%; Score 12; DB 6; Length 22;
Best Local Similarity 100.0%; Pred. No. 2.4e+04;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GCCACCATGGCC 12
Db 3 GCCACCATGGCC 14

RESULT 4
AR027668
LOCUS      AR027668      27 bp      DNA
DEFINITION Sequence 8 from patent US 5856302.
ACCESSION  AR027668
VERSION     AR027668.1 GI:5938488
KEYWORDS   .
SOURCE      Unknown.
ORGANISM    Unclassified.
REFERENCE   1 (bases 1 to 27)
AUTHORS     Ammons,W.Steve and Little,R.G.
TITLE       Therapeutic uses of bactericidal/permeability-increasing protein
            dimer products
JOURNAL     Patent: US 5856302-A 8 05-JAN-1999;
            Location/Qualifiers
            source
            1..27
            /organism="unknown"

BASE COUNT      5 a      10 c      9 g      3 t
ORIGIN

Query Match      100.0%; Score 12; DB 6; Length 27;
Best Local Similarity 100.0%; Pred. No. 2.3e+04;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GCCACCATGGCC 12
Db 10 GCCACCATGGCC 21

RESULT 5
AR050743
LOCUS      AR050743      27 bp      DNA
DEFINITION Sequence 8 from patent US 5827816.
ACCESSION  AR050743
VERSION     AR050743.1 GI:5973468
KEYWORDS   .
SOURCE      Unknown.
ORGANISM    Unclassified.
REFERENCE   1 (bases 1 to 27)
AUTHORS     Theofan,G., Horwitz,A., Burke,D., Baltaian,M. and Grinna,L.
TITLE       Stable bactericidal/permeability-increasing protein products and
            pharmaceutical compositions containing the same
JOURNAL     Patent: US 5827816-A 8 27-OCT-1998;
            Location/Qualifiers
            source
            1..27
            /organism="unknown"

BASE COUNT      5 a      10 c      9 g      3 t
ORIGIN

Query Match      100.0%; Score 12; DB 6; Length 27;
Best Local Similarity 100.0%; Pred. No. 2.3e+04;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GCCACCATGGCC 12
Db 10 GCCACCATGGCC 21

RESULT 6
AX030652
```

LOCUS AX030652 27 bp DNA PAT 20-SEP-2000
DEFINITION Sequence 8 from Patent EP1013760.
ACCESSION AX030652
VERSION AX030652.1 GI:10278178
KEYWORDS
SOURCE unidentified.
ORGANISM unidentified.
REFERENCE 1 (bases 1 to 27)
AUTHORS Burke,D., Grinna,L., Beltaian,M., Horwitz,A. and Theofan,G.
TITLE Stable bactericidal/permeability-increasing protein products and
pharmaceutical compositions containing the same
JOURNAL Patent: EP 1013760-A 8 28-JUN-2000;
XOMA TECHNOLOGY LT (BM)
FEATURES Location/Qualifiers
source 1..27
/organism="unidentified"
/db_xref="taxon:32644" 3 t
BASE COUNT 5 a 10 c 9 g 3 t
ORIGIN

Query Match 100.0%; Score 12; DB 6; Length 27;
Best Local Similarity 100.0%; Pred. No. 2.3e+04;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GCCACCATGGCC 12
|||||
Db 10 GCCACCATGGCC 21

RESULT 7
LOCUS I13573 27 bp DNA PAT 26-SEP-1995
DEFINITION Sequence 17 from patent US 5439807.
ACCESSION I13573
VERSION I13573.1 GI:996640
KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.
REFERENCE 1 (bases 1 to 27)
AUTHORS Grinna,L.S.
TITLE Methods for the preparation of endotoxin-binding proteins
JOURNAL Patent: US 5439807-A 17 08-AUG-1995;
FEATURES Location/Qualifiers
source 1..27
/organism="unknown" 3 t
BASE COUNT 5 a 10 c 9 g 3 t
ORIGIN

Query Match 100.0%; Score 12; DB 6; Length 27;
Best Local Similarity 100.0%; Pred. No. 2.3e+04;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GCCACCATGGCC 12
|||||
Db 10 GCCACCATGGCC 21

RESULT 8
LOCUS I14327 27 bp DNA PAT 26-SEP-1995
DEFINITION Sequence 8 from patent US 5447913.
ACCESSION I14327
VERSION I14327.1 GI:997342
KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.
REFERENCE 1 (bases 1 to 27)
AUTHORS Ammons,W.S. and Little,R.G.

TITLE Therapeutic uses of bactericidal/permeability-increasing protein
dimer products
JOURNAL Patent: US 5447913-A 8 05-SEP-1995;
FEATURES Location/Qualifiers
source 1..27
/organism="unknown" 3 t
BASE COUNT 5 a 10 c 9 g 3 t
ORIGIN

Query Match 100.0%; Score 12; DB 6; Length 27;
Best Local Similarity 100.0%; Pred. No. 2.3e+04;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GCCACCATGGCC 12
|||||
Db 10 GCCACCATGGCC 21

RESULT 9
LOCUS I50683 27 bp DNA PAT 07-OCT-1997
DEFINITION Sequence 17 from patent US 5643570.
ACCESSION I50683
VERSION I50683.1 GI:2472386
KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.
REFERENCE 1 (bases 1 to 27)
AUTHORS Theofan,G., Grinna,L.S. and Horwitz,A.
TITLE BpI-immunoglobulin fusion proteins
JOURNAL Patent: US 5643570-A 17 01-JUL-1997;
FEATURES Location/Qualifiers
source 1..27
/organism="unknown" 3 t
BASE COUNT 5 a 10 c 9 g 3 t
ORIGIN

Query Match 100.0%; Score 12; DB 6; Length 27;
Best Local Similarity 100.0%; Pred. No. 2.3e+04;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GCCACCATGGCC 12
|||||
Db 10 GCCACCATGGCC 21

RESULT 10
LOCUS I68187 27 bp DNA PAT 04-FEB-1998
DEFINITION Sequence 8 from patent US 5674834.
ACCESSION I68187
VERSION I68187.1 GI:2830309
KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.
REFERENCE 1 (bases 1 to 27)
AUTHORS Theofan,G., Horwitz,A., Burke,D., Beltaian,M. and Grinna,L.
TITLE Stable bactericidal/permeability-increasing protein products and
pharmaceutical compositions containing the same
JOURNAL Patent: US 5674834-A 8 07-OCT-1997;
FEATURES Location/Qualifiers
source 1..27
/organism="unknown" 3 t
BASE COUNT 5 a 10 c 9 g 3 t
ORIGIN

Query Match 100.0%; Score 12; DB 6; Length 27;
Best Local Similarity 100.0%; Pred. No. 2.3e+04;

Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GCCACCATGGCC 12
|||||
Db 10 GCCACCATGGCC 21

RESULT 11
187071
LOCUS 187071 27 bp DNA 10-JUN-1998
DEFINITION Sequence 8 from patent US 5703038.
ACCESSION I87071
VERSION I87071.1 GI:3206789
KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.
REFERENCE 1 (bases 1 to 27)
AUTHORS Ammons,W.Steve and Little,R.G.
TITLE Therapeutic uses of bactericidal-permeability-increasing protein dimer products
JOURNAL Patent: US 5703038-A 8 30-DEC-1997;
FEATURES Location/Qualifiers
source
BASE COUNT 5 a 10 c 9 g 3 t
ORIGIN

Query Match 100.0%; Score 12; DB 6; Length 27;
Best Local Similarity 100.0%; Pred. No. 2.3e+04;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GCCACCATGGCC 12
|||||
Db 10 GCCACCATGGCC 21

RESULT 12
162302
LOCUS 162302 28 bp DNA 07-OCT-1997
DEFINITION Sequence 13 from patent US 5658784.
ACCESSION I62302
VERSION I62302.1 GI:2480250
KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.
REFERENCE 1 (bases 1 to 28)
AUTHORS Eckner,R., Ewen,M. and Livingston,D.
TITLE Nucleic acid encoding transcription factor p300 and uses of p300
JOURNAL Patent: US 5658784-A 13 19-AUG-1997;
FEATURES Location/Qualifiers
source
BASE COUNT 5 a 13 c 6 g 4 t
ORIGIN

Query Match 100.0%; Score 12; DB 6; Length 28;
Best Local Similarity 100.0%; Pred. No. 2.3e+04;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GCCACCATGGCC 12
|||||
Db 11 GCCACCATGGCC 22

RESULT 13
AX179323
LOCUS AX179323 30 bp DNA 03-JUL-2001
DEFINITION Sequence 24 from Patent W00127277.
ACCESSION AX179323

VERSION AX179323.1 GI:14598994
KEYWORDS synthetic construct.
SOURCE synthetic construct.
ORGANISM artificial sequence.
REFERENCE 1 (bases 1 to 30)
AUTHORS Shimkets,R.A., Lichenstein,H. and Boldog,F.I.
TITLE Proteins and polynucleotides encoded thereby
JOURNAL Patent: WO 0127277-A 24 19-APR-2001;
FEATURES Location/Qualifiers
source
1..30
/organism="synthetic construct"
/db_xref="taxon:32630"
/note="Primer 10354784 Rev"
BASE COUNT 5 a 13 c 9 g 3 t
ORIGIN

Query Match 100.0%; Score 12; DB 6; Length 30;
Best Local Similarity 100.0%; Pred. No. 2.3e+04;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GCCACCATGGCC 12
|||||
Db 11 GCCACCATGGCC 22

RESULT 14
ARI46549
LOCUS ARI46549 32 bp DNA 08-AUG-2001
DEFINITION Sequence 9 from patent US 6218514.
ACCESSION ARI46549
VERSION ARI46549.1 GI:15109738
KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.
REFERENCE 1 (bases 1 to 32)
AUTHORS Trikha,M. and Honn,K.V.
TITLE Antibodies specific for soluble truncated integrins
JOURNAL Patent: US 6218514-A 9 17-APR-2001;
FEATURES Location/Qualifiers
source
1..32
/organism="unknown"
BASE COUNT 6 a 10 c 9 g 7 t
ORIGIN

Query Match 100.0%; Score 12; DB 6; Length 32;
Best Local Similarity 100.0%; Pred. No. 2.2e+04;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GCCACCATGGCC 12
|||||
Db 10 GCCACCATGGCC 21

RESULT 15
AR034283
LOCUS AR034283 33 bp DNA 29-SEP-1999
DEFINITION Sequence 4 from patent US 5869337.
ACCESSION AR034283
VERSION AR034283.1 GI:5949888
KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.
REFERENCE 1 (bases 1 to 33)
AUTHORS Crabtree,G.R., Schreiber,S.L., Spencer,D.M., Wandless,T.J. and Belshaw,P.
TITLE Regulated transcription of targeted genes and other biological events

JOURNAL Patent: US 5869337-A 4 09-FEB-1999;

FEATURES

Location/Qualifiers

Source

1. .33

/organism="unknown"

BASE COUNT 8 a 13 c 9 g 3 t

ORIGIN

Query Match

Best Local Similarity 100.0%; Score 12; DB 6; Length 33;

Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GCCACCATGGCC 12

Db 11 GCCACCATGGCC 22

Search completed: January 17, 2002, 11:49:54

Job time: 17206 sec

THIS PAGE BLANK (USPTO)

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: January 17, 2002, 11:52:06 ; Search time 221.34 Seconds
(without alignments)
11.255 Million cell updates/sec

Title: US-09-242-202a-10_COPY_1_11
Perfect score: 11
Sequence: 1 GCCTTAAGGC 11

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 351203 seqs, 113238999 residues

Total number of hits satisfying chosen parameters: 702406

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued_Patents_NA.*
1: /cgn2_6/ptodata/2/ina/5A_COMB.seq.*
2: /cgn2_6/ptodata/2/ina/5B_COMB.seq.*
3: /cgn2_6/ptodata/2/ina/6A_COMB.seq.*
4: /cgn2_6/ptodata/2/ina/6B_COMB.seq.*
5: /cgn2_6/ptodata/2/ina/PCTUS_COMB.seq.*
6: /cgn2_6/ptodata/2/ina/backfiles1.seq.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	11	100.0	17	4	US-09-384-305-17
2	11	100.0	17	4	US-09-258-377-22
3	11	100.0	33	2	US-08-873-479-39
4	11	100.0	33	4	US-09-384-305-23
5	11	100.0	33	4	US-09-258-377-28
6	11	100.0	33	4	US-09-258-377-31
7	11	100.0	34	2	US-08-873-479-19
8	11	100.0	34	2	US-08-972-661A-19
c 9	11	100.0	36	2	US-08-873-479-18
c 10	11	100.0	36	2	US-08-972-661A-18
11	11	100.0	37	2	US-09-031-442A-7
12	11	100.0	37	4	US-09-258-377-7
c 13	11	100.0	40	2	US-08-873-479-46
c 14	11	100.0	40	2	US-08-972-661A-32
15	11	100.0	42	2	US-08-873-479-37
16	11	100.0	42	2	US-09-031-442A-9
17	11	100.0	42	2	US-08-972-661A-25
18	11	100.0	42	4	US-09-258-377-9
19	11	100.0	48	2	US-08-873-479-45
20	11	100.0	48	2	US-08-972-661A-31
21	11	100.0	54	2	US-09-031-442A-1
22	11	100.0	54	4	US-09-384-305-5
23	11	100.0	54	4	US-09-258-377-1
c 24	11	100.0	57	2	US-08-972-661A-28
25	11	100.0	58	2	US-09-031-442A-3
26	11	100.0	58	2	US-08-972-661A-27
27	11	100.0	58	4	US-09-384-305-7

28	11	100.0	58	4	US-09-258-377-3	Sequence 3, Appl
29	11	100.0	185	4	US-09-384-305-20	Sequence 20, Appl
30	11	100.0	185	4	US-09-384-305-21	Sequence 21, Appl
31	11	100.0	185	4	US-09-384-305-22	Sequence 22, Appl
32	11	100.0	185	4	US-09-258-377-25	Sequence 25, Appl
33	11	100.0	185	4	US-09-258-377-26	Sequence 26, Appl
34	11	100.0	185	4	US-09-258-377-27	Sequence 27, Appl
c 35	11	100.0	687	3	US-09-048-052-5	Sequence 5, Appl
c 36	11	100.0	687	4	US-09-160-246-5	Sequence 5, Appl
37	11	100.0	860	1	US-08-310-416A-19	Sequence 19, Appl
38	11	100.0	860	2	US-08-888-171-19	Sequence 19, Appl
39	11	100.0	969	1	US-08-310-416A-12	Sequence 12, Appl
40	11	100.0	969	2	US-08-888-171-12	Sequence 12, Appl
c 41	11	100.0	985	4	US-09-160-246-9	Sequence 9, Appl
c 42	11	100.0	985	4	US-09-160-246-13	Sequence 13, Appl
43	11	100.0	1426	4	US-09-230-380-4	Sequence 4, Appl
44	11	100.0	1878	2	US-08-435-149-17	Sequence 17, Appl
45	11	100.0	1886	6	5256558-1	Patent No. 5256558

ALIGNMENTS

RESULT 1
US-09-384-305-17
; Sequence 17, Application US/09384305
; Patent No. 6184028
; GENERAL INFORMATION:
; APPLICANT: Michael D. Thomas
; TITLE OF INVENTION: Polypeptides Having Pectin
; TITLE OF INVENTION: Acetyltransferase Activity And Nucleic Acids Encoding Same
; FILE REFERENCE: 5952,000-US
; CURRENT APPLICATION NUMBER: US/09/384,305
; CURRENT FILING DATE: 1999-08-26
; NUMBER OF SEQ ID NOS: 25
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 17
; LENGTH: 17
; TYPE: DNA
; ORGANISM: Bacillus subtilis
US-09-384-305-17

Query Match 100.0%; Score 11; DB 4; Length 17;
Best Local Similarity 100.0%; Pred. No. 80;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GCCTTAAGGC 11
|||||||
Db 2 gccttaagggc 12

RESULT 2
US-09-258-377-22
; Sequence 22, Application US/09258377
; Patent No. 6255076
; GENERAL INFORMATION:
; APPLICANT: Widner, William
; APPLICANT: Sloma, Michael D.
; APPLICANT: Thomas, Michael D.
; TITLE OF INVENTION: Methods For Producing A polypeptide In a
; TITLE OF INVENTION: Bacillus Cell
; FILE REFERENCE: 5455,200-US
; CURRENT APPLICATION NUMBER: US/09/258,377
; CURRENT FILING DATE: 1999-02-26
; EARLIER FILING DATE: 1999-02-26
; EARLIER FILING DATE: 1998-02-26
; NUMBER OF SEQ ID NOS: 33
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 22
; LENGTH: 17
; TYPE: DNA

; TITLE OF INVENTION: Acetyltransferase Activity And Nucleic Acids Encoding Same
; FILE REFERENCE: 5952,000-US
; CURRENT APPLICATION NUMBER: US/09/384,305
; CURRENT FILING DATE: 1999-08-26
; NUMBER OF SEQ ID NOS: 25
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 23
; LENGTH: 33
; TYPE: DNA
; ORGANISM: Bacillus subtilis
; US-09-384-305-23

Query Match 100.0%; Score 11; DB 4; Length 33;
Best Local Similarity 100.0%; Pred. No. 85;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 GCCTTAAGGCG 11
Db 2 gccttaaggcg 12
|||||

RESULT 5
US-09-258-377-28
; Sequence 28, Application US/09258377
; Patent No. 6255076
; GENERAL INFORMATION:
; APPLICANT: Widner, William
; APPLICANT: Sloma, Alan
; APPLICANT: Thomas, Michael D.
; TITLE OF INVENTION: Methods For Producing A polypeptide In a
; TITLE OF INVENTION: Bacillus Cell
; FILE REFERENCE: 5455,200-US
; CURRENT APPLICATION NUMBER: US/09/258,377
; CURRENT FILING DATE: 1999-02-26
; EARLIER FILING DATE: 1998-02-26
; EARLIER FILING DATE: 1998-02-26
; NUMBER OF SEQ ID NOS: 33
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 28
; LENGTH: 33
; TYPE: DNA
; ORGANISM: Bacillus
; US-09-258-377-28

Query Match 100.0%; Score 11; DB 4; Length 33;
Best Local Similarity 100.0%; Pred. No. 85;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 GCCTTAAGGCG 11
Db 2 gccttaaggcg 12
|||||

RESULT 6
US-09-258-377-31
; Sequence 31, Application US/09258377
; Patent No. 6255076
; GENERAL INFORMATION:
; APPLICANT: Widner, William
; APPLICANT: Sloma, Alan
; APPLICANT: Thomas, Michael D.
; TITLE OF INVENTION: Methods For Producing A polypeptide In a
; TITLE OF INVENTION: Bacillus Cell
; FILE REFERENCE: 5455,200-US
; CURRENT APPLICATION NUMBER: US/09/258,377
; CURRENT FILING DATE: 1999-02-26
; EARLIER FILING DATE: 1998-02-26
; EARLIER FILING DATE: 1998-02-26
; NUMBER OF SEQ ID NOS: 33
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 31

; ORGANISM: Bacillus
US-09-258-377-22
Query Match 100.0%; Score 11; DB 4; Length 17;
Best Local Similarity 100.0%; Pred. No. 80;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 GCCTTAAGGCG 11
Db 2 gccttaaggcg 12
|||||

RESULT 3
US-08-873-479-39
; Sequence 39, Application US/08873479
; Patent No. 5891701
; GENERAL INFORMATION:
; APPLICANT: Sloma, Alan
; APPLICANT: Lynne, Christianson
; TITLE OF INVENTION: Nucleic Acids Encoding A Polypeptide
; TITLE OF INVENTION: Having Protease Activity
; NUMBER OF SEQUENCES: 57
; CORRESPONDENCE ADDRESSES:
; ADDRESS: No. 5891701 No. 5891701disk of No. 5891701th America
; STREET: 405 Lexington Avenue
; CITY: New York
; STATE: NY
; COUNTRY: USA
; ZIP: 10174
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/873,479
; FILING DATE: 12-JUN-1997
; CLASSIFICATION: 530
; ATTORNEY/AGENT INFORMATION:
; NAME: Agtis, Cheryl H
; REGISTRATION NUMBER: 34,086
; REFERENCE/DOCKET NUMBER: 5251,000-US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 212-867-0123
; TELEFAX: 212-878-9655
; TELEX:
; INFORMATION FOR SEQ ID NO: 39:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 33 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
US-08-873-479-39

Query Match 100.0%; Score 11; DB 2; Length 33;
Best Local Similarity 100.0%; Pred. No. 85;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 GCCTTAAGGCG 11
Db 5 GCCTTAAGGCG 15
|||||

RESULT 4
US-09-384-305-23
; Sequence 23, Application US/09384305
; Patent No. 6184028
; GENERAL INFORMATION:
; APPLICANT: Michael D. Thomas
; APPLICANT: Kimberly M. Brown
; TITLE OF INVENTION: Polypeptides Having Pectin


```
; LENGTH: 33
; TYPE: DNA
; ORGANISM: Bacillus
US-09-258-377-31

Query Match      100.0%; Score 11; DB 4; Length 33;
Best Local Similarity 100.0%; Pred. No. 85;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GCCTTAAGGC 11
Db 2 gccttaaggc 12

RESULT 7
US-08-873-479-19
; Sequence 19, Application US/08873479
; Patent No. 5891701
; GENERAL INFORMATION:
; APPLICANT: Sloma, Alan
; TITLE OF INVENTION: Nucleic Acids Encoding A Polypeptide
; NUMBER OF SEQUENCES: 57
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: No. 5891701o No. 5891701disk of No. 5891701th America
; STREET: 405 Lexington Avenue
; CITY: New York
; STATE: NY
; COUNTRY: USA
; ZIP: 10174
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/873,479
; FILING DATE: 12-JUN-1997
; CLASSIFICATION: 530
; ATTORNEY/AGENT INFORMATION:
; NAME: Agris, Cheryl H
; REGISTRATION NUMBER: 34,086
; REFERENCE/DOCKET NUMBER: 5251.000-US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 212-867-0123
; TELEFAX: 212-878-9655
; TELEX:
; INFORMATION FOR SEQ ID NO: 19:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 34 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
US-08-873-479-19

Query Match      100.0%; Score 11; DB 2; Length 34;
Best Local Similarity 100.0%; Pred. No. 85;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GCCTTAAGGC 11
Db 3 GCCTTAAGGC 13

RESULT 8
US-08-972-661A-19
; Sequence 19, Application US/08872661A
; Patent No. 5958728
; GENERAL INFORMATION:
; APPLICANT: Sloma, Alan
```

```
; APPLICANT: Sternberg, David
; APPLICANT: Adams, Lee F.
; APPLICANT: Brown, Stephen
; TITLE OF INVENTION: Method For Producing Polypeptides
; TITLE OF INVENTION: In Mutants Of Bacillus Cells
; NUMBER OF SEQUENCES: 35
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: No. 5958728o No. 5958728disk Of No. 5958728th America, Inc.
; STREET: 405 Lexington Avenue - 64th Fl.
; CITY: New York
; STATE: NY
; COUNTRY: USA
; ZIP: 10174
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/972,661A
; FILING DATE: 18-NOV-1997
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Staines, Robert L
; REGISTRATION NUMBER: 41,324
; REFERENCE/DOCKET NUMBER: 5111.200-US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 212-878-9652
; TELEFAX: 212-878-9655
; INFORMATION FOR SEQ ID NO: 19:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 34 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
US-08-972-661A-19

Query Match      100.0%; Score 11; DB 2; Length 34;
Best Local Similarity 100.0%; Pred. No. 85;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GCCTTAAGGC 11
Db 3 GCCTTAAGGC 13

RESULT 9
US-08-873-479-18/c
; Sequence 18, Application US/08873479
; Patent No. 5891701
; GENERAL INFORMATION:
; APPLICANT: Sloma, Alan
; APPLICANT: Lynne, Christianson
; TITLE OF INVENTION: Nucleic Acids Encoding A Polypeptide
; TITLE OF INVENTION: Having Protease Activity
; NUMBER OF SEQUENCES: 57
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: No. 5891701o No. 5891701disk of No. 5891701th America
; STREET: 405 Lexington Avenue
; CITY: New York
; STATE: NY
; COUNTRY: USA
; ZIP: 10174
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/873,479
; FILING DATE: 12-JUN-1997
; CLASSIFICATION: 530
; ATTORNEY/AGENT INFORMATION:
; NAME: Agris, Cheryl H
; REGISTRATION NUMBER: 34,086
; REFERENCE/DOCKET NUMBER: 5251.000-US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 212-867-0123
; TELEFAX: 212-878-9655
; TELEX:
; INFORMATION FOR SEQ ID NO: 19:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 34 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
US-08-873-479-19
```

ATTORNEY/AGENT INFORMATION:
NAME: Agilis, Cheryl H
REGISTRATION NUMBER: 34,086
REFERENCE/DOCKET NUMBER: 5251.000-US
TELECOMMUNICATION INFORMATION:
TELEPHONE: 212-867-0123
TELEFAX: 212-878-9655
TELEX:
INFORMATION FOR SEQ ID NO: 18:
SEQUENCE CHARACTERISTICS:
LENGTH: 36 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
US-08-873-479-18

Query Match 100.0%; Score 11; DB 2; Length 36;
Best Local Similarity 100.0%; Pred. No. 85;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GCCTTAAGGCG 11
Db 22 GCCTTAAGGCG 12

RESULT 10
US-08-972-661A-18/C
Sequence 18, Application US/08972661A
Patent No. 5958728
GENERAL INFORMATION:
APPLICANT: Sloma, Alan
APPLICANT: Sternberg, David
APPLICANT: Adams, Lee F.
APPLICANT: Brown, Stephen
TITLE OF INVENTION: Method For Producing Polypeptides
TITLE OF INVENTION: In Mutants Of Bacillus Cells
NUMBER OF SEQUENCES: 35
CORRESPONDENCE ADDRESS:
ADDRESSEE: No. 5958728 Of No. 5958728th America, Inc.
STREET: 405 Lexington Avenue - 64th Fl.
CITY: New York
STATE: NY
COUNTRY: USA
ZIP: 10174
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: FastSEQ for Windows Version 2.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/972,661A
FILING DATE: 18-NOV-1997
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Starnes, Robert L
REGISTRATION NUMBER: 41,324
REFERENCE/DOCKET NUMBER: 5111.200-US
TELECOMMUNICATION INFORMATION:
TELEPHONE: 212-878-9652
TELEFAX: 212-878-9655
INFORMATION FOR SEQ ID NO: 18:
SEQUENCE CHARACTERISTICS:
LENGTH: 36 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
US-08-972-661A-18

Query Match 100.0%; Score 11; DB 2; Length 36;
Best Local Similarity 100.0%; Pred. No. 85;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GCCTTAAGGCG 11
Db 22 GCCTTAAGGCG 12

RESULT 11
US-09-031-442A-7
Sequence 7, Application US/09031442A
Patent No. 5955310
GENERAL INFORMATION:
APPLICANT: Widner, William
APPLICANT: Sloma, Alan
APPLICANT: Thomas, Michael D.
TITLE OF INVENTION: Methods For Producing A Polypeptide
TITLE OF INVENTION: In A Bacillus Cell
NUMBER OF SEQUENCES: 21
CORRESPONDENCE ADDRESS:
ADDRESSEE: No. 5955310 No. 5955310th America, Inc.
STREET: 405 Lexington Avenue
CITY: New York
STATE: NY
COUNTRY: U.S.A.
ZIP: 10174
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: FastSEQ for Windows Version 2.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/031,442A
FILING DATE: 26-FEB-1998
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Lambiris, Elias J.
REGISTRATION NUMBER: 33,728
REFERENCE/DOCKET NUMBER: 5455.000-US
TELECOMMUNICATION INFORMATION:
TELEPHONE: 212-867-0123
TELEFAX: 212-878-9655
TELEX:
INFORMATION FOR SEQ ID NO: 7:
SEQUENCE CHARACTERISTICS:
LENGTH: 37 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
US-09-031-442A-7

Query Match 100.0%; Score 11; DB 2; Length 37;
Best Local Similarity 100.0%; Pred. No. 85;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GCCTTAAGGCG 11
Db 5 GCCTTAAGGCG 15

RESULT 12
US-09-258-377-7
Sequence 7, Application US/09258377
Patent No. 6255076
GENERAL INFORMATION:
APPLICANT: Widner, William
APPLICANT: Sloma, Alan
APPLICANT: Thomas, Michael D.
TITLE OF INVENTION: Methods For Producing A polypeptide In a
TITLE OF INVENTION: Bacillus Cell
FILE REFERENCE: 5455.200-US
CURRENT APPLICATION NUMBER: US/09/258,377
CURRENT FILING DATE: 1999-02-26
EARLIER APPLICATION NUMBER: 09/031,442

; EARLIER FILING DATE: 1998-02-26
 ; NUMBER OF SEQ ID NOS: 33
 ; SOFTWARE: FastSEQ for Windows Version 3.0
 ; SEQ ID NO: 7
 ; LENGTH: 37
 ; TYPE: DNA
 ; ORGANISM: Bacillus
 ; US-09-258-377-7

Query Match 100.0%; Score 11; DB 4; Length 37;
 Best Local Similarity 100.0%; Pred. No. 85;
 Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GCCTTAAGGC 11
 DB 5 gccttaaggcc 15

RESULT 13
 US-08-873-479-46/C
 ; Sequence 46, Application US/08873479
 ; Patent No. 5891701
 ; GENERAL INFORMATION:
 ; APPLICANT: Sloma, Alan
 ; APPLICANT: Lynne, Christianson
 ; TITLE OF INVENTION: Nucleic Acids Encoding A Polypeptide
 ; TITLE OF INVENTION: Having Protease Activity
 ; NUMBER OF SEQUENCES: 57
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: No. 5891701o No. 5891701disk of No. 5891701th America
 ; STREET: 405 Lexington Avenue
 ; CITY: New York
 ; STATE: NY
 ; COUNTRY: USA
 ; ZIP: 10174
 ; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: Diskette
 ; COMPUTER: IBM Compatible
 ; OPERATING SYSTEM: DOS
 ; SOFTWARE: FastSEQ for Windows Version 2.0
 ; CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: US/08/873,479
 ; FILING DATE: 12-JUN-1997
 ; CLASSIFICATION: 530
 ; ATTORNEY/AGENT INFORMATION:
 ; NAME: Agis, Cheryl H
 ; REGISTRATION NUMBER: 34,086
 ; REFERENCE/DOCKET NUMBER: 5251.000-US
 ; TELECOMMUNICATION INFORMATION:
 ; TELEPHONE: 212-867-0123
 ; TELEFAX: 212-878-9655
 ; TELEX:

; INFORMATION FOR SEQ ID NO: 46:
 ; SEQUENCE CHARACTERISTICS:
 ; LENGTH: 40 base pairs
 ; TYPE: nucleic acid
 ; STRANDEDNESS: single
 ; TOPOLOGY: linear
 ; US-08-873-479-46

Query Match 100.0%; Score 11; DB 2; Length 40;
 Best Local Similarity 100.0%; Pred. No. 86;
 Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GCCTTAAGGC 11
 DB 38 GCCTTAAGGC 28

RESULT 14
 US-08-972-661A-32/C

; Sequence 32, Application US/08972661A
 ; Patent No. 5958728
 ; GENERAL INFORMATION:
 ; APPLICANT: Sloma, Alan
 ; APPLICANT: Sternberg, David
 ; APPLICANT: Adams, Lee F.
 ; APPLICANT: Brown, Stephen
 ; TITLE OF INVENTION: Method For Producing Polypeptides
 ; TITLE OF INVENTION: In Mutants Of Bacillus Cells
 ; NUMBER OF SEQUENCES: 35
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: No. 5958728o No. 5958728disk of No. 5958728th America, Inc.
 ; STREET: 405 Lexington Avenue - 64th Fl.
 ; CITY: New York
 ; STATE: NY
 ; COUNTRY: USA
 ; ZIP: 10174
 ; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: Diskette
 ; COMPUTER: IBM Compatible
 ; OPERATING SYSTEM: DOS
 ; SOFTWARE: FastSEQ for Windows Version 2.0
 ; CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: US/08/972,661A
 ; FILING DATE: 18-NOV-1997
 ; CLASSIFICATION: 435
 ; ATTORNEY/AGENT INFORMATION:
 ; NAME: Starnes, Robert L
 ; REGISTRATION NUMBER: 41,324
 ; REFERENCE/DOCKET NUMBER: 5111.200-US
 ; TELECOMMUNICATION INFORMATION:
 ; TELEPHONE: 212-878-9652
 ; TELEFAX: 212-878-9655
 ; INFORMATION FOR SEQ ID NO: 32:
 ; SEQUENCE CHARACTERISTICS:
 ; LENGTH: 40 base pairs
 ; TYPE: nucleic acid
 ; STRANDEDNESS: single
 ; TOPOLOGY: linear
 ; US-08-972-661A-32

Query Match 100.0%; Score 11; DB 2; Length 40;
 Best Local Similarity 100.0%; Pred. No. 86;
 Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GCCTTAAGGC 11
 DB 38 GCCTTAAGGC 28

RESULT 15
 US-08-873-479-37
 ; Sequence 37, Application US/08873479
 ; Patent No. 5891701
 ; GENERAL INFORMATION:
 ; APPLICANT: Sloma, Alan
 ; APPLICANT: Lynne, Christianson
 ; TITLE OF INVENTION: Nucleic Acids Encoding A Polypeptide
 ; TITLE OF INVENTION: Having Protease Activity
 ; NUMBER OF SEQUENCES: 57
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: No. 5891701o No. 5891701disk of No. 5891701th America
 ; STREET: 405 Lexington Avenue
 ; CITY: New York
 ; STATE: NY
 ; COUNTRY: USA
 ; ZIP: 10174
 ; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: Diskette
 ; COMPUTER: IBM Compatible
 ; OPERATING SYSTEM: DOS
 ; SOFTWARE: FastSEQ for Windows Version 2.0

```

; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/873,479
; FILING DATE: 12-JUN-1997
; CLASSIFICATION: 530
; ATTORNEY/AGENT INFORMATION:
; NAME: Agiris, Cheryl H
; REGISTRATION NUMBER: 34,086
; REFERENCE/DOCKET NUMBER: 5251.000-US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 212-867-0123
; TELEFAX: 212-878-9655
; TELEX:
; INFORMATION FOR SEQ ID NO: 37:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 42 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
;
US-08-873-479-37

```

```

Query Match      100.0%; Score 11; DB 2; Length 42;
Best Local Similarity 100.0%; Pred. No. 86;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 GCCTTAAGGC 11
        |||||
Db       5 GCCTTAAGGC 15

```

Search completed: January 17, 2002, 11:52:06
Job time: 17253 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: January 17, 2002, 07:03:08 ; Search time 10436.3 Seconds
(without alignments)
358.829 Million cell updates/sec

Title: US-09-242-202a-10

Perfect score: 227

Sequence: 1 GCCTTAGGCCATATGGTG.....GGTGTCTCTATATATATAT 227

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 1472140 seqs, 8248589755 residues

Total number of hits satisfying chosen parameters: 2944280

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

GenEmbl:*

1: gb_ba:*

2: gb_htg:*

3: gb_in:*

4: gb_om:*

5: gb_ov:*

6: gb_pat:*

7: gb_ph:*

8: gb_pl:*

9: gb_pr:*

10: gb_ro:*

11: gb_sts:*

12: gb_sy:*

13: gb_un:*

14: gb_vi:*

15: em_ba:*

16: em_fun:*

17: em_in:*

18: em_in:*

19: em_om:*

20: em_or:*

21: em_ov:*

22: em_pat:*

23: em_ph:*

24: em_pl:*

25: em_ro:*

26: em_sts:*

27: em_sy:*

28: em_un:*

29: em_vi:*

30: em_htgo_hum:*

31: em_htgo_inv:*

32: em_htgo_rod:*

33: em_htgo_hum:*

34: em_htg_inv:*

35: em_htg_rod:*

36: em_htg_other:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No. Score Match Length DB ID Description

1	105.4	46.4	1964	9	HSGROW2	V00520 Human germ
2	105.4	46.4	2771	6	I41411	I41411 Sequence 5
3	105.4	46.4	4456	6	AR119909	AR119909 Sequence
4	105.4	46.4	4874	12	AF239248	AF239248 Eukaryoti
5	98	43.2	1827	9	AF237583	AF237583 Homo sapi
6	96	42.3	831	6	I02857	I02857 Sequence 3
7	96	42.3	2086	6	AR102885	AR102885 Sequence
8	96	42.3	2086	6	AR108064	AR108064 Sequence
9	96	42.3	2086	6	AR134712	AR134712 Sequence
10	96	42.3	2657	9	HUMGHN	M13438 Human growt
11	96	42.3	3141	6	A18662	A18662 Synthetic c
12	96	42.3	3141	6	A18663	A18663 Synthetic c
13	96	42.3	3141	6	A20119	A20119 Hybrid insu
14	96	42.3	3141	6	A20120	A20120 Hybrid insu
15	96	42.3	4495	12	AF369966	AF369966 Cloning v
16	96	42.3	4650	12	AF396260	AF396260 Cloning v
17	96	42.3	4657	12	AF239249	AF239249 Eukaryoti
18	96	42.3	4665	12	AF239250	AF239250 Eukaryoti
19	96	42.3	4707	12	AF067196	AF067196 Cloning v
20	96	42.3	4713	12	AF067197	AF067197 Cloning v
21	96	42.3	4724	12	AF239247	AF239247 Eukaryoti
22	96	42.3	4775	12	AF239251	AF239251 Eukaryoti
23	96	42.3	4912	12	AF239252	AF239252 Eukaryoti
24	96	42.3	5646	9	HS47670	AJ007670 Homo sapi
25	96	42.3	9274	6	AX060703	AX060703 Sequence
26	96	42.3	66495	9	HUMGHCSA	J03071 Human growt
27	96	42.3	159128	2	AC040958	AC040958 Homo sapi
28	96	42.3	159387	2	AC084858	AC084858 Homo sapi
29	95.6	42.1	1944	9	AF237585	AF237585 Homo sapi
30	95	41.9	6776	12	PVPHM3	X76682 Plasmid vec
31	95	41.9	8451	12	PVPHM3BL	X76683 Plasmid vec
32	87	38.3	2660	6	E00140	E00140 Genomic DNA
33	87	38.3	2660	9	HUMGHV	K00470 Homo sapien
34	87	38.3	2740	9	HUMCS3	M15894 Human chori
35	87	38.3	159128	2	AC040958	AC040958 Homo sapi
36	87	38.3	191583	9	AC015651	AC015651 Homo sapi
37	86.4	38.1	1125	6	I02858	I02858 Sequence 4
38	86.4	38.1	2033	9	AF237586	AF237586 Homo sapi
39	86.4	38.1	2967	9	HUMPLA	J00289 Homo sapien
40	85.4	37.6	2594	9	MM002293	U02293 Macaca mula
41	75	33.0	2193	9	AF237584	AF237584 Homo sapi
42	74.2	32.7	6160	1	ECT7DN3	Z32693 E.coli pT7h
43	74.2	32.7	11271	1	ECT7DN1	Z32691 E.coli (HMI
44	74	32.6	642	9	AF110644	AF110644 Homo sapi
45	74	32.6	796	6	E00009	E00009 mRNA coding

ALIGNMENTS

RESULT 1

HSGROW2

LOCUS

DEFINITION

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

MEDLINE

AUTHORS

TITLE

HSGROW2 1964 bp DNA PRI 10-FEB-1999
Human germ line gene for growth hormone (presomatotropin).

V00520 J00148 K00612

V00520.1 GI:31906

complementary DNA; germ line; growth hormone; signal peptide.
human.

Homo sapiens

Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;

Mammalia; Euthera; Primates; Catarrhini; Hominidae; Homo.

1 (bases 1 to 1964)

DeNoto,F.M., Moore,D.D. and Goodman,H.M.

Human growth hormone DNA sequence and mRNA structure: possible

alternative splicing

Nucleic Acids Res. 9 (15), 3719-3730 (1981)

82014939

2 (bases 1 to 1964)

Adelman,J.P., Hayflick,J.S., Vasser,M. and Seeburg,P.H.

In vitro deletional mutagenesis for bacterial production of the

JOURNAL DNA 2 (3), 183-193 (1983)
MEDLINE 84057143
COMMENT This entry was previously called <HSGROWL>.
See <HSGROWL> for mRNA sequence (with some differences).
FEATURES Location/Qualifiers

source 1..1964
/organism="Homo sapiens"
/db_xref="taxon:9606"
prim_transcript (274,275)..1905
mRNA join(274,275)..344,601..761,971..1090,1184..1348,1602..1905)
exon (274,275)..344
/number=1
CDS join(335..344,601..761,971..1090,1184..1348,1602..1799)
/codon_start=1
/product="growth hormone"
/protein_id="CAA23779.1"
/db_xref="GI:312406"
/translation="MATGSETSLIAFLCLLPWLGSAFPTIPLSLFDNASLRAH
RLHQAEDTTQEEFETIPEKQYKFLQNTSLCFSESIPTSNRETOQKSNLELL
RISLLLSWLEPWFQFVRSFANSLVYGASDNYDLKDLDEGIQTLMGRLDGSPR
TGQIFKQYTKFEDTNSHNDALLKNYLLYCFRKMDKVETFLRIVQCRSVEGSGF"
345..600
intron /number=1
exon 601..761
/number=2
intron 762..970
/number=2
exon 971..1090
/number=3
intron 1091..1183
/number=3
exon 1184..1348
/number=4
intron 1349..1601
/number=4
exon 1602..1905
/number=5

BASE COUNT 451 a 550 g 413 t
ORIGIN
Query Match 46.4%; Score 105.4; DB 9; Length 1964;
Best Local Similarity 99.1%; Pred. No. 7.6e-21;
Matches 106; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 121 AGTGCCTCTCTCGGCGCTGGAAGTTGCCACTCCAGTGGCCACCGCCTTGCTCTAATAAA 180
|||||
Db 1832 AGTGCCTCTCTCGGCGCTGGAAGTTGCCACTCCAGTGGCCACCGCCTTGCTCTAATAAA 1891
|||||
QY 181 ATTAAGTTGCATCATTTTGTCTGACTAGGTGTCTCTATATATATAT 227
|||||
Db 1892 ATTAAGTTGCATCATTTTGTCTGACTAGGTGTCTCTATATATATAT 1938
|||||
RESULT 2
I41411
LOCUS I41411 2771 bp DNA PAT 13-MAY-1997
DEFINITION Sequence 5 from patent US 5625124.
ACCESSION I41411
VERSION I41411.1 GI:2082001
KEYWORDS
SOURCE Unknown.
ORGANISM Unclassified.
REFERENCE 1 (bases 1 to 2771):
AUTHORS Falk, P. and Gordon, J. I.
TITLE Animal model for helicobacter pylori infection
JOURNAL Patent: US 5625124-A 5 29-APR-1997;
FEATURES Location/Qualifiers
source 1..2771

BASE COUNT 666 a 718 c 704 g 683 t
ORIGIN
Query Match 46.4%; Score 105.4; DB 6; Length 2771;
Best Local Similarity 99.1%; Pred. No. 7.5e-21;
Matches 106; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 121 AGTGCCTCTCTCGGCGCTGGAAGTTGCCACTCCAGTGGCCACCGCCTTGCTCTAATAAA 180
|||||
Db 2177 AGTGCCTCTCTCGGCGCTGGAAGTTGCCACTCCAGTGGCCACCGCCTTGCTCTAATAAA 2236
|||||
QY 181 ATTAAGTTGCATCATTTTGTCTGACTAGGTGTCTCTATATATATAT 227
|||||
Db 2237 ATTAAGTTGCATCATTTTGTCTGACTAGGTGTCTCTATATATATAT 2283
|||||
RESULT 3
AR119909
LOCUS AR119909 4456 bp DNA PAT 16-MAY-2001
DEFINITION Sequence 1 from patent US 6153427.
ACCESSION AR119909
VERSION AR119909.1 GI:14102608
KEYWORDS
SOURCE Unknown.
ORGANISM Unclassified.
REFERENCE 1 (bases 1 to 4456)
AUTHORS King, D. and Schneider, G. B.
TITLE Erythropoietin-inducible, erythroid-specific DNA construct
JOURNAL Patent: US 6153427-A 1 28-NOV-2000;
FEATURES Location/Qualifiers
source 1..4456
BASE COUNT 1238 a 978 c 1076 g 1164 t
ORIGIN

Query Match 46.4%; Score 105.4; DB 6; Length 4456;
Best Local Similarity 99.1%; Pred. No. 7.3e-21;
Matches 106; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 121 AGTGCCTCTCTCGGCGCTGGAAGTTGCCACTCCAGTGGCCACCGCCTTGCTCTAATAAA 180
|||||
Db 4324 AGTGCCTCTCTCGGCGCTGGAAGTTGCCACTCCAGTGGCCACCGCCTTGCTCTAATAAA 4383
|||||
QY 181 ATTAAGTTGCATCATTTTGTCTGACTAGGTGTCTCTATATATATAT 227
|||||
Db 4384 ATTAAGTTGCATCATTTTGTCTGACTAGGTGTCTCTATATATATAT 4430
|||||
RESULT 4
AF239248
LOCUS AF239248 4874 bp DNA circular SYN 12-APR-2000
DEFINITION Eukaryotic expression vector pCMV4, complete sequence.
ACCESSION AF239248
VERSION AF239248.1 GI:7542545
KEYWORDS
SOURCE Expression vector pCMV4.
ORGANISM Expression vector pCMV4.
REFERENCE 1 (bases 1 to 4874)
AUTHORS Andersson, S., Davis, D.L., Dahlback, H., Jornvall, H. and Russell, D.W.
TITLE Cloning, structure, and expression of the mitochondrial cytochrome
P-450 sterol 26-hydroxylase, a bile acid biosynthetic enzyme
JOURNAL J. Biol. Chem. 264 (14), 8222-8229 (1989)
MEDLINE 89255259
PUBMED 2722778
REFERENCE 2 (bases 1 to 4874)
AUTHORS Andersson, S., Davis, D.L., Dahlback, H., Jornvall, H. and Russell, D.W.
TITLE Direct Submission
JOURNAL Submitted (25-FEB-2000) Molecular Genetics, University of Texas

Query Match 46.4%; Score 105.4; DB 9; Length 1964;
Best Local Similarity 99.1%; Pred. No. 7.6e-21;
Matches 106; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 121 AGTGCCTCTCTCGGCGCTGGAAGTTGCCACTCCAGTGGCCACCGCCTTGCTCTAATAAA 180
|||||
Db 1832 AGTGCCTCTCTCGGCGCTGGAAGTTGCCACTCCAGTGGCCACCGCCTTGCTCTAATAAA 1891
|||||
QY 181 ATTAAGTTGCATCATTTTGTCTGACTAGGTGTCTCTATATATATAT 227
|||||
Db 1892 ATTAAGTTGCATCATTTTGTCTGACTAGGTGTCTCTATATATATAT 1938
|||||
RESULT 2
I41411
LOCUS I41411 2771 bp DNA PAT 13-MAY-1997
DEFINITION Sequence 5 from patent US 5625124.
ACCESSION I41411
VERSION I41411.1 GI:2082001
KEYWORDS
SOURCE Unknown.
ORGANISM Unclassified.
REFERENCE 1 (bases 1 to 2771):
AUTHORS Falk, P. and Gordon, J. I.
TITLE Animal model for helicobacter pylori infection
JOURNAL Patent: US 5625124-A 5 29-APR-1997;
FEATURES Location/Qualifiers
source 1..2771

Southwestern Medical Center at Dallas, 5323 Harry Hines Blvd.,
Dallas, TX 75235, USA

FEATURES Location/Qualifiers

1..4874

/organism="Expression vector pCMV4"

/db_xref="taxon:120119"

BASE COUNT 1165 a 1237 c 1170 g 1302 t

ORIGIN

Query Match 46.4%; Score 105.4; DB 12; Length 4874;
Best Local Similarity 99.1%; Pred. No. 7.3e-21;
Matches 106; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 121 AGTGCTCTCTGCGCCCTGGAAGTTGCCACTCCAGTGCCCGCCAGCCCTTGCCCTAATAAA 180

|||||

Db 1055 AGTGCTCTCTGCGCCCTGGAAGTTGCCACTCCAGTGCCCGCCAGCCCTTGCCCTAATAAA 1114

|||||

QY 181 ATTAAGTTGCATCATTTTGTCTGACTAGGTGTCCTCTATAATATAT 227

|||||

Db 1115 ATTAAGTTGCATCATTTTGTCTGACTAGGTGTCCTCTATAATATAT 1161

RESULT 5
AF237583 1827 bp DNA PRI 11-MAY-2001
LOCUS Homo sapiens recombinant IgG1 heavy chain gene, partial cds.
DEFINITION
ACCESSION AF237583
VERSION AF237583.1 GI:9857752
KEYWORDS
SOURCE human.

ORGANISM
REFERENCE 1 (bases 1 to 1827)
AUTHORS Vidarsson, G., Jansen, M., Boel, E. and van de Winkel, J.G.J.
TITLE Activity of Human IgG and IgA Subclasses in Immune Defense Against
Neisseria meningitidis Serogroup B

J. Immunol. 166 (10), 6250-6256 (2001)
11342648
2 (bases 1 to 1827)
AUTHORS Vidarsson, G., Jansen, M., Boel, E. and van de Winkel, J.G.J.
TITLE Direct Submission
JOURNAL Submitted (22-FEB-2000) Department of Immunology, University
Medical Center Utrecht, Rm. KC.02-085.2, Lundlaan 6, Utrecht 3584
EA, The Netherlands

FEATURES Location/Qualifiers
source 1..1827
exon 1..294
mRNA join(<1..294,686..730,849..1178,1276..>1598)
CDS join(<1..294,686..730,849..1178,1276..1598)
/codon_start=3
/product="recombinant IgG1 heavy chain"
/protein_id="AAG00909.1"
/db_xref="GI:9857753"

/translation="STKGPSVFPLAPSSKSTSGGTAAALGCLVKDYFPEPTVSWNSGA
LTSVHPPFAPVLOSGLYSIVSVVTPFSSSLGTQTYICNVNHKFSNTKVDKRVPEKSC
DKRTHCPAPFELGGPSVFLPPKPKDTLMISRTPEVTCVVVDVSHEDPEVKFNWY
VDGVEHNKATPREEDYNTSRVYSLVTLVHODWLNKEYKKVSKNALPAPIEKTI
SKAKGQPREPVYTLPSREMTKNQVSLTCLVKGFPYSDIAVWEWSNGALPENNKKIT
PPVLDSGGSFELYSLKLVDRKSRWQQGVFSCVMHEALHNHYTQKSLSLSPGK"

686..730
/note="hinge"
849..1178
/note="CH2"
1276..1598
/note="CH3"

BASE COUNT 390 a 621 c 498 g 316 t 2 others
ORIGIN

Query Match 43.2%; Score 98; DB 9; Length 1827;
Best Local Similarity 78.3%; Pred. No. 1.2e-18;
Matches 130; Conservative 0; Mismatches 35; Indels 1; Gaps 1;

QY 63 TCAGAGCCCCGGGCGGACGACAGGCCAATGCCGCTCTTCCCTCGAGGATGAGTAGTGAG 122

|||||

Db 1578 TCCCTGTCCCCGGGTAATAGTGGGTACCGGGTGGCATCCCTGTGCCCTCCCCAG 1637

|||||

QY 123 TGCCCTCTCTGCGCCCTGGAAGTTGCCACTCCAGTGCCCGCCAGCCCTTGCCCTAATAAAAT 182

|||||

Db 1638 TGCCCTCTCTGCGCCCTGGAAGTTGCCACTCCAGTGCCCGCCAGCCCTTGCCCTAATAAAAT 1697

|||||

QY 183 TAAAGTTGCATCATTTTGTCTGACTAGGTGTCCTCTATAATATAT 227

|||||

Db 1698 TAAAGTTGCATCATTTTGTCTGACTAGGTGTCCTCTATAATATAT 1743

|||||

RESULT 6
I02857 831 bp ss-DNA PAT 18-MAY-1993
LOCUS Sequence 3 from Patent US 4446235.
DEFINITION
ACCESSION I02857
VERSION I02857.1 GI:267919
KEYWORDS
SOURCE Unknown.
ORGANISM Unclassified.

REFERENCE 1 (bases 1 to 831)
AUTHORS Seeburg, P.H.
TITLE Method for cloning human growth hormone variant genes
JOURNAL Patent: US 4446235-A 3 01-MAY-1984;
Genentech, Inc.;
South San Francisco, CA

FEATURES Location/Qualifiers
source 1..831
BASE COUNT 175 a 231 c 219 g 206 t
ORIGIN

Query Match 42.3%; Score 96; DB 6; Length 831;
Best Local Similarity 99.1%; Pred. No. 5e-18;
Matches 107; Conservative 0; Mismatches 0; Indels 1; Gaps 1;

QY 121 AGTGCTCTCTGCGCCCTGGAAGTTGCCACTCCAGTGCCCGCCAGCCCTTGCTCTAATAAA 180

|||||

Db 231 AGTGCTCTCTGCGCCCTGGAAGTTGCCACTCCAGTGCCCGCCAGCCCTTGCTCTAATAAA 290

|||||

QY 181 ATTAAGTTGCATCATTTTGTCTGACTAGGTGTCCTCTATAATATAT 227

|||||

Db 291 ATTAAGTTGCATCATTTTGTCTGACTAGGTGTCCTCTATAATATAT 338

|||||

RESULT 7
AR102885 2086 bp DNA PAT 14-FEB-2001
LOCUS Sequence 9 from patent US 6087129.
DEFINITION
ACCESSION AR102885
VERSION AR102885.1 GI:12814473
KEYWORDS
SOURCE Unknown.
ORGANISM Unclassified.

REFERENCE 1 (bases 1 to 2086)
AUTHORS Newgard, C.B., Normington, K.D., Clark, S.A., Thigpen, A.E., Quaaide, C.
and Kruse, F.
TITLE Recombinant expression of proteins from secretory cell lines
JOURNAL Patent: US 6087129-A 9 11-JUL-2000;
FEATURES Location/Qualifiers

DEFINITION Synthetic cDNA sequence for hybrid insulin/IGF-I receptor.
ACCESSION A18663
VERSION A18663.1 GI:512264

KEYWORDS
SOURCE synthetic construct.
ORGANISM synthetic construct
artificial sequence.

REFERENCE
1 (bases 1 to 3141)

AUTHORS
TITLE A HYBRID CELLULAR RECEPTOR

JOURNAL
Patent: WO 9117253-A 15 14-NOV-1991;

FEATURES
source
1. .3141
/organism="synthetic construct"

/db_xref="taxon:32630"

BASE COUNT 669 a 863 c 892 g 717 t

ORIGIN

Query Match 42.3%; Score 96; DB 6; Length 3141;
Best Local Similarity 99.1%; Pred. No. 4.7e-18;

Matches 107; Conservative 0; Mismatches 0; Indels 1; Gaps 1;

QY 121 AGTGCCTCTCTCGCCCTGGAGTTGCCACTCCAGTCCGCCACCCAGCCTTGCTCTAATAAA 180
|||||

Db 244 AGTGCCTCTCTCGCCCTGGAGTTGCCACTCCAGTCCGCCACCCAGCCTTGCTCTAATAAA 185
|||||

QY 181 ATTAAGTTGCATCATTTTGTCTGACTAGGTGTCCTCTATAATATTAT 227
|||||

Db 184 ATTAAGTTGCATCATTTTGTCTGACTAGGTGTCCTCTATAATATTAT 137
|||||

RESULT 13

A20119

LOCUS Hybrid insulin/IGF-I receptor mRNA

DEFINITION 3141 bp PAT 08-AUG-1994

ACCESSION A20119

VERSION A20119.1 GI:583286

KEYWORDS
SOURCE synthetic construct.
ORGANISM synthetic construct
artificial sequence.

REFERENCE
1 (bases 1 to 3141)

AUTHORS
TITLE A HYBRID CELLULAR RECEPTOR

JOURNAL
Patent: WO 9117252-A 15 14-NOV-1991;

FEATURES
source
1. .3141
/organism="synthetic construct"

/db_xref="taxon:32630"

BASE COUNT 717 a 892 c 863 g 669 t

ORIGIN

Query Match 42.3%; Score 96; DB 6; Length 3141;
Best Local Similarity 99.1%; Pred. No. 4.7e-18;

Matches 107; Conservative 0; Mismatches 0; Indels 1; Gaps 1;

QY 121 AGTGCCTCTCTCGCCCTGGAGTTGCCACTCCAGTCCGCCACCCAGCCTTGCTCTAATAAA 180
|||||

Db 2898 AGTGCCTCTCTCGCCCTGGAGTTGCCACTCCAGTCCGCCACCCAGCCTTGCTCTAATAAA 2957
|||||

QY 181 ATTAAGTTGCATCATTTTGTCTGACTAGGTGTCCTCTATAATATTAT 227
|||||

Db 2958 ATTAAGTTGCATCATTTTGTCTGACTAGGTGTCCTCTATAATATTAT 3005
|||||

RESULT 14

A20120/c

LOCUS Hybrid insulin/IGF-I receptor mRNA

DEFINITION 3141 bp PAT 08-AUG-1994

ACCESSION A20120

VERSION A20120.1 GI:578998

KEYWORDS

SOURCE synthetic construct.
ORGANISM synthetic construct
artificial sequence.

REFERENCE
1 (bases 1 to 3141)

AUTHORS
TITLE A HYBRID CELLULAR RECEPTOR

JOURNAL
Patent: WO 9117252-A 16 14-NOV-1991;

FEATURES
source
1. .3141
/organism="synthetic construct"

/db_xref="taxon:32630"

BASE COUNT 669 a 863 c 892 g 717 t

ORIGIN

Query Match 42.3%; Score 96; DB 6; Length 3141;
Best Local Similarity 99.1%; Pred. No. 4.7e-18;

Matches 107; Conservative 0; Mismatches 0; Indels 1; Gaps 1;

QY 121 AGTGCCTCTCTCGCCCTGGAGTTGCCACTCCAGTCCGCCACCCAGCCTTGCTCTAATAAA 180
|||||

Db 244 AGTGCCTCTCTCGCCCTGGAGTTGCCACTCCAGTCCGCCACCCAGCCTTGCTCTAATAAA 185
|||||

QY 181 ATTAAGTTGCATCATTTTGTCTGACTAGGTGTCCTCTATAATATTAT 227
|||||

Db 184 ATTAAGTTGCATCATTTTGTCTGACTAGGTGTCCTCTATAATATTAT 137
|||||

RESULT 15

AF369966

LOCUS

DEFINITION 4495 bp DNA circular SYN 22-JUL-2001

ACCESSION AF369966

VERSION AF369966.1 GI:14994098

KEYWORDS
SOURCE Cloning vector pCMV-MCS.
ORGANISM Cloning vector pCMV-MCS.
artificial sequence; vectors.

REFERENCE
1 (bases 1 to 4495)

AUTHORS
TITLE Grafsky,A.J. III.

JOURNAL
pCMV-MCS for use in the AAV Helper-Free System

REFERENCE
2 (bases 1 to 4495)

AUTHORS
TITLE Grafsky,A.J. III.

JOURNAL
Direct Submission

FEATURES
Submitted (16-Apr-2001) Technical Services, Stratagene, 11011 N.
Torrey Pines Rd., La Jolla, CA 92037, USA

Location/Qualifiers
1. .4495
/organism="Cloning vector pCMV-MCS"

/db_xref="taxon:161367"

/lab_host="Escherichia coli"

promoter
1. .670
/note="CMV promoter"

misc_feature
678. .1170
/note="similar to beta-globin intron"

misc_feature
1178. .1253
/note="MCS"

misc_feature
1249. .1735
/note="hGH/poly A"

rep_origin
complement(1844. .2701)
/note="ColE1 Origin"

CDS
complement(2711. .3571)
/note="confers ampicillin resistance"

/codon_start=1
/transl_table=11
/product="beta-lactamase"

/protein_id="AAK76424.1"

/db_xref="GI:14994098"

/translation="MSIOHFRVALIPFAAFCLPFAHETLVKVKDAEDQIGARVGY
IELDLSNGKILSFSPREPEFPMMSTFKVLLCGNVLSDRDAQEQQLGRRHYSONDLVE
YSPVTEKHLTDGNTVRELCSAITSMDNTAANLLLTITIGPKELTAFLNMGDHVTPL

THIS PAGE BLANK (USPTO)

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OW nucleic - nucleic search, using sw model

Run on: January 17, 2002, 07:06:53 ; Search time 599.86 Seconds
(without alignments)
324.430 Million cell updates/sec

Title: US-09-242-202a-10
Perfect score: 227
Sequence: 1 GCCTTAAGGCCATATGTTG.....GGTGTCTCTATAATATTAT 227

Scoring table: IDENTITY NUC
Gapop 10.0 , Gapext 1.0

Searched: 930621 seqs, 428662619 residues

Total number of hits satisfying chosen parameters: 1861242

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : N_Geneseq_1101.*

- 1: /SIDS2/gcgdata/geneseq/geneseq/NA1980.DAT.*
- 2: /SIDS2/gcgdata/geneseq/geneseq/NA1981.DAT.*
- 3: /SIDS2/gcgdata/geneseq/geneseq/NA1982.DAT.*
- 4: /SIDS2/gcgdata/geneseq/geneseq/NA1983.DAT.*
- 5: /SIDS2/gcgdata/geneseq/geneseq/NA1984.DAT.*
- 6: /SIDS2/gcgdata/geneseq/geneseq/NA1985.DAT.*
- 7: /SIDS2/gcgdata/geneseq/geneseq/NA1986.DAT.*
- 8: /SIDS2/gcgdata/geneseq/geneseq/NA1987.DAT.*
- 9: /SIDS2/gcgdata/geneseq/geneseq/NA1988.DAT.*
- 10: /SIDS2/gcgdata/geneseq/geneseq/NA1989.DAT.*
- 11: /SIDS2/gcgdata/geneseq/geneseq/NA1990.DAT.*
- 12: /SIDS2/gcgdata/geneseq/geneseq/NA1991.DAT.*
- 13: /SIDS2/gcgdata/geneseq/geneseq/NA1992.DAT.*
- 14: /SIDS2/gcgdata/geneseq/geneseq/NA1993.DAT.*
- 15: /SIDS2/gcgdata/geneseq/geneseq/NA1994.DAT.*
- 16: /SIDS2/gcgdata/geneseq/geneseq/NA1995.DAT.*
- 17: /SIDS2/gcgdata/geneseq/geneseq/NA1996.DAT.*
- 18: /SIDS2/gcgdata/geneseq/geneseq/NA1997.DAT.*
- 19: /SIDS2/gcgdata/geneseq/geneseq/NA1998.DAT.*
- 20: /SIDS2/gcgdata/geneseq/geneseq/NA1999.DAT.*
- 21: /SIDS2/gcgdata/geneseq/geneseq/NA2000.DAT.*
- 22: /SIDS2/gcgdata/geneseq/geneseq/NA2001.DAT.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	227	100.0	227	19 AAV21719	Combined 3' splice
2	214.4	94.4	228	19 AAV21723	Humanised vector s
3	214.4	94.4	1547	19 AAV21732	Humanised vector p
4	214.4	94.4	1807	19 AAV21733	Humanised vector p
5	214.4	94.4	2308	19 AAV21734	Humanised vector p
6	160.6	70.7	1425	19 AAV21724	Humanised vector p
7	160.6	70.7	2125	19 AAV21727	Humanised vector p
8	152	67.0	1911	19 AAV21726	Humanised vector p
9	105.4	46.4	2771	18 AAV76770	Rat Fabpl gut-spec
10	105.4	46.4	4456	22 AAC83268	Nucleotide sequenc
11	96	42.3	191	19 AAV33343	Human growth hormo

12	96	42.3	191	20 AA240410	Human growth hormo
13	96	42.3	191	21 AA250393	Human growth hormo
14	96	42.3	563	22 AAF76860	Human secreted pro
15	96	42.3	1441	20 AAX15886	Not1/Sali psk-GHRH
16	96	42.3	2086	21 AAC55720	Secretory cell DNA
17	96	42.3	2086	22 AAF58805	Human growth hormo
18	96	42.3	2160	18 AAT62826	Human growth hormo
19	96	42.3	3141	12 AAQ14381	Extracellular doma
20	96	42.3	3141	12 AAQ14813	Soluble human insu
21	96	42.3	3141	12 AAQ14815	Hybrid human insul
22	96	42.3	3426	20 AA240416	Plasmid pIF0836 co
23	96	42.3	3589	20 AA240418	Plasmid pIF0921 co
24	96	42.3	3599	19 AAV50428	Plasmid pIG0552 lo
25	96	42.3	3599	19 AAV40796	Actual sequence of
26	96	42.3	3600	19 AAV50427	Plasmid pIG0552 up
27	96	42.3	3600	19 AAV40795	Expected sequence
28	96	42.3	5686	20 AA240415	Plasmid pIN1143 co
29	96	42.3	6513	21 AA475627	Nucleotide sequenc
30	96	42.3	9164	20 AAX82259	Beta-domain delete
31	96	42.3	9274	22 AAF24297	HER2 transgene pla
32	96	42.3	9534	18 AAT62072	Vector RP3224E2 en
33	96	42.3	11846	20 AAX82261	Factor VIII protei
34	96	42.3	12022	20 AAX82260	Factor VIII protei
35	95	41.9	44576	21 AA261522	Cosmid CV014 conta
36	93.4	41.1	3653	18 AAT59270	Plasmid pMPTVE4ORF
37	87	38.3	2660	4 AAN30032	Sequence of gene f
38	79	34.8	2994	22 AAF84109	Human cancer speci
39	74	32.6	682	15 AAT02620	Met-des(F1-P2-P3-I
40	74	32.6	682	16 AAT11228	Met-des(F1-P2-P3-I
41	74	32.6	700	16 AAT02619	Met-des(Phe1-Pro2-
42	74	32.6	700	16 AAT11227	Met-des(Phe1-Pro2-
43	74	32.6	706	16 AAT02618	Met-des(Phe1-Pro2-
44	74	32.6	706	16 AAT11225	Met-des(Phe1-Pro2)
45	74	32.6	709	16 AAT02617	Met-des(Phe1)-soma

ALIGNMENTS

RESULT 1
AAV21719
ID AAV21719 standard; cDNA; 227 BP.
XX
AC AAV21719;
XX
DT 17-AUG-1998 (first entry)
XX
DE Combined 3' splice sequence and polyA tail.
XX
KW Vector; vaccine; tumour; antigen; ds.
OS Homo sapiens.
OS Synthetic.
XX
PN WO9806863-Al.
XX
PD 19-FEB-1998.
XX
PF 14-AUG-1997; 97WO-US14306.
XX
PR 14-AUG-1996; 96US-0023931.
XX
PA (USSH) US DEPT HEALTH & HUMAN SERVICES.
XX
PI Nelson EL, Nelson PJ;
XX
DR WPI; 1998-159552/14.
XX
PT Humanised polynucleotide vectors - comprising human derived promoter
XX and sequence acceptance site, used for the production of vaccines
PS Disclosure; Page 17; 125pp; English.
XX

CC This DNA sequence comprises a combined 3' splice sequence and
 CC polyA tail sequence derived from human gene sequences. Novel
 CC humanised vectors of the invention (see AAV21724, AAV21727 and
 CC AAV21732-34) comprise a human-derived promoter or mammalian homologue
 CC which is functional in mammalian target tissue and cells and an
 CC acceptance site which accepts cDNA products from RT-PCR cloning.
 CC They may also include human-derived splice-polyA sequences. The
 CC vectors are used to express target antigens, especially tumour
 CC antigens. They are non-replicating in mammalian cells but are
 CC capable of extended stable expression of target sequences generating
 CC an immune response in immunised individuals. The vectors
 CC selectively elicit immune responses to the target sequences with
 CC little or no immune response to the other components of the vectors.
 CC The target antigens are expressed as intracellular polypeptides or
 CC peptides and, as such, are processed as self polypeptides or
 CC peptides and appropriately presented on antigen presenting cells.
 XX
 SQ Sequence 227 BP; 47 A; 64 C; 60 G; 56 T; 0 other;

Query Match 100.0%; Score 227; DB 19; Length 227;
 Best Local Similarity 100.0%; Pred. No. 1.4e-62;
 Matches 227; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 GCCTTAAGGGCCATATGGTGGATGCCCTTGACCCAGCCGGGATGGGAGACCTGT 60
 DB 1 GCCTTAAGGGCCATATGGTGGATGCCCTTGACCCAGCCGGGATGGGAGACCTGT 60
 QY 61 AGTCAGAGCCCCGGGAGCAGCAGCCAAATGCCGCTCTCCCTGCAGGATGAGTAGTG 120
 DB 61 AGTCAGAGCCCCGGGAGCAGCAGCCAAATGCCGCTCTCCCTGCAGGATGAGTAGTG 120
 QY 121 AGTCGCTCTCTCTGCGCCCTGGGAAGTTGGCAGTCCAGTCCACCCAGCCCTTGTCCTAATAA 180
 DB 121 AGTCGCTCTCTGCGCCCTGGGAAGTTGGCAGTCCAGTCCACCCAGCCCTTGTCCTAATAA 180
 QY 181 ATTAAGTTGCATATTTGCTGACTAGGTGCTCTCTAATAATATAT 227
 DB 181 ATTAAGTTGCATATTTGCTGACTAGGTGCTCTCTAATAATATAT 227

RESULT 2
 AAV21723
 ID AAV21723 standard; cDNA; 228 BP.
 AC AAV21723;
 XX 17-AUG-1998 (first entry)
 DE Humanised vector splice polyA signal sequence.
 KW Vector; vaccine; tumour; antigen; ds.
 XX Homo sapiens.
 OS Synthetic.
 FH Key Location/Qualifiers
 FT misc_feature 1..17
 FT /*tag= a
 FT /note= "potential site for IRES"
 FT intron 18..120
 FT /*tag= b
 FT /number= 3
 FT exon 121..228
 FT /*tag= c
 FT /number= 4
 FT polyA_signal 176..181
 FT /*tag= d
 XX WO9806863-A1.
 XX 19-FEB-1998.
 PD
 XX

PF 14-AUG-1997; 97WO-US14306.
 XX
 PR 14-AUG-1996; 96US-0023931.
 XX
 PA (USSH) US DEPT HEALTH & HUMAN SERVICES.
 XX
 XI Nelson EL, Nelson PJ;
 XX WPI; 1998-159552/14.
 DR
 XX Humanised polynucleotide vectors - comprising human derived promoter
 PT and sequence acceptance site, used for the production of vaccines
 XX
 PS Example 1; Page 27; 125pp; English.
 XX
 CC This DNA sequence comprises a combined 3' splice sequence and
 CC polyA tail sequence cloned into the Xho and HindIII sites of
 CC pGEM-7zf. Novel humanised vectors of the invention (see AAV21724,
 CC AAV21727 and AAV21732-34) comprise a human-derived promoter or mammalian
 CC homologue which is functional in mammalian target tissue and cells
 CC and an acceptance site which accepts cDNA products from RT-PCR
 CC cloning. They may also include the splice-polyA signal
 CC sequence. After processing of mRNA, intron 3 is excised, and
 CC the sequence contains duplicate stop codons in 2 reading frames.
 CC A 3rd reading frame stop codon can be obtained by mutagenising
 CC the internal ribosomal entry site (IRES). The vectors are used to
 CC express target antigens, especially tumour antigens. They are
 CC non-replicating in mammalian cells but are capable of extended
 CC stable expression of target sequences, and selectively elicit
 CC immune responses to the target sequences with little or no immune
 CC response to the other vector components. The target antigens are
 CC expressed as intracellular polypeptides and are appropriately
 CC presented on antigen presenting cells.
 XX
 SQ Sequence 228 BP; 47 A; 63 C; 62 G; 56 T; 0 other;

Query Match 94.4%; Score 214.4; DB 19; Length 228;
 Best Local Similarity 99.1%; Pred. No. 1.3e-58;
 Matches 226; Conservative 0; Mismatches 1; Indels 1; Gaps 1;
 QY 1 GCCTTAAGGGCCATATGGTGGATGCCCTTGACCCAGCCGGGAT-GGGAGACCTG 59
 DB 1 GCCTTAAGGGCCATATGGTGGATGCCCTTGACCCAGCCGGGATGgggggagacctg 60
 QY 60 TAGTCAGAGCCCCGGGAGCAGCAGCCAAATGCCGCTCTCCCTGCAGGATGAGTAGT 119
 DB 61 tagtcagagccccgggagcagcagccaaatgccgctctctccctgcaggatgagtagt 120
 QY 120 GAGTGCCTCTCTGCGCCCTGGGAAGTTGCCACTCCAGTCCACCCAGCCCTTGTCCTAATAA 179
 DB 121 gagtgcctctctgccccgggagtgccactccagtgccccagcctgtgcctaataa 180
 QY 180 AATTAAGTTGCATATTTGCTGACTAGGTGCTCTCTAATAATATAT 227
 DB 181 aattaagttgcataatttctgactaggtgctctctataataattat 228

RESULT 3
 AAV21732
 ID AAV21732 standard; cDNA; 1547 BP.
 AC AAV21732;
 XX 17-AUG-1998 (first entry)
 DE Humanised vector pITL-A.
 KW Vector; vaccine; tumour; antigen; plasmid pITL-A; ds.
 XX Chimeric - Homo sapiens.
 OS Chimeric - Escherichia coli.
 OS Synthetic.

QY 1 GCCTTAAGGCCATATGGTGGATGCCCTTGACCCAGCGGGGAT-GGGGAGACCTG 59
 Db 232 gccttaagggccatattggtgagtgagtccttgacccagcgaggatgggggagacctg 291
 QY 60 TAGTCAGAGCCCCGGGAGCAGCACAGGCCAATGCCGCTCTTCCCTGCAGGATGAGTAGT 119
 Db 292 tagtcagagccccggcgagcacagggccaaagccgctctccctgcaggatgagtagt 351
 QY 120 GAGTGCCCTCTCTGCGCCCTGGAAGTTGCCACTCCAGTCCCGCAGCCAGCCTTCTCTTAATAA 179
 Db 352 gagtgcctctcctggcctggaagtggccactccagtgcccaccagcctgtctctaataa 411
 QY 180 AATTAAGTTGCATATTTGTCTGACTAGGTGCTCTATATAATATAT 227
 Db 412 aattaagttgcatatttgtctgactagggtgtctctataattat 459

RESULT 5
 AAV21734
 ID AAV21734 standard; cDNA; 2308 BP.
 XX AC AAV21734;
 XX 17-AUG-1998 (first entry)
 XX Humanised vector pITL-1 GFP.
 DE Vector; vaccine; tumour; antigen; plasmid pITL-1 GFP;
 KW green fluorescent protein; ds.
 KW Chimeric - Homo sapiens.
 OS Chimeric - Escherichia coli.
 OS Chimeric - Aequorea victoria.
 OS Synthetic.

Key Location/Qualifiers
 FT CDS 20..734
 FT /*tag= a
 FT /product= green fluorescent protein
 FT 973..1181
 FT /*tag= c
 FT /note= "SupF gene"
 FT 1191..1793
 FT /*tag= d
 FT /note= "ColE1 origin of replication"
 FT 2063..2308
 FT /*tag= e
 FT /note= "RANTES promoter"

WO9806863-A1.
 XX 19-FEB-1998.
 XX 14-AUG-1997; 97WO-US14306.
 XX 14-AUG-1996; 96US-0023931.
 XX (USSH) US DEPT HEALTH & HUMAN SERVICES.
 PA Nelson EL, Nelson PJ;
 PI WPI; 1998-159552/14.
 XX Humanised polynucleotide vectors - comprising human derived promoter
 FT and sequence acceptance site, used for the production of vaccines
 XX Example 11; Page 57-58; 125pp; English.
 CC Plasmid pITL-1 GFP comprises base vector pITL-1 (see AAV21733) and
 CC a humanised green fluorescent protein (GFP) reporter sequence
 CC (see AAV21725). Novel humanised vectors comprise a human-derived
 CC promoter or mammalian homologue which is functional in mammalian
 CC target tissue and cells and a sequence acceptance site which

CC accepts cDNA products from RT-PCR cloning. The novel vectors are
 CC used to express target antigens, especially tumour antigens. They
 CC are non-replicating in mammalian cells but are capable of extended
 CC stable expression of target sequences generating an immune response
 CC in immunised individuals. The vectors selectively elicit immune
 CC responses to the target sequences with little or no immune response
 CC to the other components of the vectors. The target antigens are
 CC expressed as intracellular polypeptides or peptides and, as such,
 CC are processed as self polypeptides or peptides and appropriately
 CC presented on antigen presenting cells.
 XX Sequence 2308 BP; 571 A; 601 C; 604 G; 532 T; 0 other;
 SQ

Query Match 94.4%; Score 214.4; DB 19; Length 2308;
 Best Local Similarity 99.1%; Pred. No. 3.3e-58;
 Matches 226; Conservative 0; Mismatches 1; Indels 1; Gaps 1;

QY 1 GCCTTAAGGCCATATGGTGGATGCCCTTGACCCAGCGGGGAT-GGGGAGACCTG 59
 Db 733 gccttaagggccatattggtgagtgagtccttgacccagcgaggatgggggagacctg 792
 QY 60 TAGTCAGAGCCCCGGGAGCAGCACAGGCCAATGCCGCTCTTCCCTGCAGGATGAGTAGT 119
 Db 793 tagtcagagccccggcgagcacagggccaaagccgctctccctgcaggatgagtagt 852
 QY 120 GAGTGCCCTCTCTGCGCCCTGGAAGTTGCCACTCCAGTCCCGCAGCCAGCCTTCTCTTAATAA 179
 Db 853 gagtgcctctcctggcctggaagtggccactccagtgcccaccagcctgtctctaataa 912
 QY 180 AATTAAGTTGCATATTTGTCTGACTAGGTGCTCTATATAATATAT 227
 Db 913 aattaagttgcatatttgtctgactagggtgtctctataattat 960

RESULT 6
 AAV21724
 ID AAV21724 standard; cDNA; 1425 BP.
 XX AC AAV21724;
 XX 17-AUG-1998 (first entry)
 XX Humanised vector pITL.
 DE Vector; vaccine; tumour; antigen; plasmid pITL; ds.
 KW Chimeric - Homo sapiens.
 OS Chimeric - Escherichia coli.
 OS Synthetic.

Key Location/Qualifiers
 FT misc_feature 1..221
 FT /*tag= a
 FT /note= "stuffer sequence"
 FT 22..481
 FT /*tag= b
 FT /note= "combined splice and polyA sequences"
 FT 495..701
 FT /*tag= c
 FT /note= "SupF gene"
 FT 712..1164
 FT /*tag= d
 FT /note= "ColE1 origin of replication"
 FT 1177..1425
 FT /*tag= e
 FT /note= "RANTES promoter"

WO9806863-A1.
 XX 19-FEB-1998.
 XX 14-AUG-1997; 97WO-US14306.

XX 14-AUG-1996; 96US-0023931.
 XX (USSH) US DEPT HEALTH & HUMAN SERVICES.
 XX Nelson EL, Nelson PJ;
 XX WPI; 1998-159552/14.
 XX Humanised polynucleotide vectors - comprising human derived promoter
 PT and sequence acceptance site, used for the production of vaccines
 XX Claim 14; Page 29-30; 125pp; English.
 XX Plasmid pITL comprises a base vector for novel humanised
 CC polynucleotide vectors. Such vectors comprise a human-derived
 CC promoter or mammalian homologue which is functional in mammalian
 CC target tissue and cells and a sequence acceptance site (see
 CC AAV21735-36), which accepts cDNA products from RT-PCR cloning. They
 CC also contain minimal non-human components, such as a replication
 CC origin (see AAV21715) and selectable marker gene (see AAV21717-18) that
 CC are necessary for production of the vector, as well as human-derived
 CC splice and polyA sequences (see AAV21723). The novel vectors are
 CC used to express target antigens, especially tumour antigens. They
 CC are non-replicating in mammalian cells but are capable of extended
 CC stable expression of target sequences generating an immune response
 CC in immunised individuals. The vectors selectively elicit immune
 CC responses to the target sequences with little or no immune response
 CC to the other components of the vectors. The target antigens are
 CC expressed as intracellular polypeptides or peptides and, as such,
 CC are processed as self polypeptides or peptides and appropriately
 CC presented on antigen presenting cells.
 XX Sequence 1425 BP; 325 A; 381 C; 404 G; 315 T; 0 other;

Query Match 70.7%; Score 160.6; DB 19; Length 1425;
 Best Local Similarity 84.2%; Pred. No. 3.1e-41;
 Matches 218; Conservative 0; Mismatches 9; Indels 32; Gaps 2;
 QY 1 GCCTTAAGGGCCATATGGTGGAGTGCCTTACCCAGCGGGGAT-GGGGAGACCTG 59
 DB 223 gctctaaggccatattggtgagtgatgcttgacccagcgggagtgaggacctg 282
 QY 60 TAGTCAGAGCCCCGGGAGCAGACGCCAATGCCGCTCTCCCTCGCAG----- 109
 DB 283 tagtcagagccccgggagcagacagcccaatgcccgtccctccctgcaagttagtga 342
 QY 110 -----GATGAGTAGTGGTGCCTCTCTCGCCCTGGAAGTTGCC 148
 DB 343 ctgcccgggtggatccctgtgacccctcccccagtgccctcctgcccctggaagtggc 402
 QY 149 ACTCAGTCCCCACAGCCTTCTCTATAATAAATTAAGTTGATCATTTTGTCTGACTAG 208
 DB 403 actccagtgccacagcctgtcctataataataaattgaattgcatctattgtcgtactag 462
 QY 209 GTGCTCTCTATAATATAT 227
 DB 463 gtgtcctctataattat 481
 RESULT 7
 AAV21727
 ID AAV21727 standard; cDNA; 2125 BP.
 XX
 AC AAV21727;
 XX
 DT 17-AUG-1998 (first entry)
 XX
 DE Humanised vector pITL-hHER/neu.
 XX
 KW Vector; vaccine; tumour; antigen; plasmid pITL-hHER/neu;
 KW human; HER-2/neu; C-erbB-2; breast cancer; ds.

XX Chimeric - Homo sapiens.
 OS Chimeric - Escherichia coli.
 OS Synthetic.
 XX Key Location/Qualifiers
 FH 13..921 /tag= a
 FT /product= human HER-2/neu
 FT 922..1181 /tag= b
 FT /note= "combined splice and polyA sequences"
 FT 1195..1401 /tag= c
 FT /note= "SupF gene"
 FT 1412..1864 /tag= d
 FT /note= "ColE1 origin of replication"
 FT 1877..2125 /tag= e
 FT /note= "RANTES promoter"
 XX WO9806863-A1.
 PN 19-FEB-1998.
 XX 14-AUG-1997; 97WO-US14306.
 XX 14-AUG-1996; 96US-0023931.
 XX (USSH) US DEPT HEALTH & HUMAN SERVICES.
 PA Nelson EL, Nelson PJ;
 XX WPI; 1998-159552/14.
 XX Humanised polynucleotide vectors - comprising human derived promoter
 PT and sequence acceptance site, used for the production of vaccines
 XX Example 7; Page 41-42; 125pp; English.
 XX Plasmid pITL-hHER2/neu comprises base vector pITL (see AAV21724) and
 CC a human HER-2/neu nucleic sequence. pITL-hHER2/neu was used to
 CC evaluate the toxicity of anti-tumour vaccination in rats, and in
 CC phase I and phase II trials to evaluate polynucleotide vaccination
 CC in advanced breast cancer. Novel humanised vectors, which can be
 CC based on pITL, comprise a human-derived promoter or mammalian
 CC homologue which is functional in mammalian target tissue and cells
 CC and a sequence acceptance site which accepts cDNA products from
 CC RT-PCR cloning. The vectors are non-replicating in mammalian cells
 CC but are capable of extended stable expression of the target
 CC sequence, generating an immune response in immunised individuals.
 CC The vectors selectively elicit immune responses to the target
 CC sequences with little or no immune response to the other components
 CC of the vectors.
 XX Sequence 2125 BP; 449 A; 650 C; 598 G; 428 T; 0 other;

Query Match 70.7%; Score 160.6; DB 19; Length 2125;
 Best Local Similarity 84.2%; Pred. No. 3.7e-41;
 Matches 218; Conservative 0; Mismatches 9; Indels 32; Gaps 2;
 QY 1 GCCTTAAGGGCCATATGGTGGAGTGCCTTACCCAGCGGGGAT-GGGGAGACCTG 59
 DB 922 gctctaaggccatattggtgagtgatgcttgacccagcgggagtgaggacctg 981
 QY 60 TAGTCAGAGCCCCGGGAGCAGACGCCAATGCCGCTCTCCCTCGCAG----- 109
 DB 982 tagtcagagccccgggagcagacagcccaatgcccgtccctccctgcaagttagtga 1041
 QY 110 -----GATGAGTAGTGGTGCCTCTCTCGCCCTGGAAGTTGCC 148
 DB 463 gtgtcctctataattat 481

Db 1042 ctgccgggtgggataccctgtgacccctccccagtgctctctgcccctgggaagttgccc 1101

QY 149 ACTCCAGTCCACCCAGCTGTGCTTAATAAATAAATTAAGTTGTCATCATTTTGTCTGACTAG 208

Db 1102 actccagtgcccacccgctgctccctaataaaaataagttgcatactttgtgtgactag 1161

QY 209 GTGCTCTCTATAATATTAT 227

Db 1162 gtgtctctataattat 1180

RESULT 8

AAV21726

ID AAV21726 standard; cDNA; 1911 BP.

XX

AC AAV21726;

XX

DT 17-AUG-1998 (first entry)

XX

XX Humanised vector pITL-GFP.

XX

KW Vector; vaccine; tumour; antigen; plasmid pITL-GFP;

KW green fluorescent protein; ds.

XX

OS Chimeric - Homo sapiens.

OS Chimeric - Escherichia coli.

OS Chimeric - Aequorea victoria.

OS Synthetic.

XX

XX Key Location/Qualifiers

FT CDS 2..719

FT FT /*tag= a

FT FT /product= green fluorescent protein

FT polyA_site 720..967

FT FT /*tag= b

FT FT /note= "combined splice and polyA sequences"

FT CDS 981..1187

FT FT /*tag= c

FT FT /note= "SupF gene"

FT misc_feature 1198..1650

FT FT /*tag= d

FT FT /note= "ColEI origin of replication"

FT promoter 1663..1911

FT FT /*tag= e

FT FT /note= "RANTES promoter"

XX

XX WO9806863-A1.

XX

XX 19-FEB-1998.

XX

XX 14-AUG-1997; 97WO-US14306.

XX

XX 14-AUG-1996; 96US-0023931.

XX

XX (USSH) US DEPT HEALTH & HUMAN SERVICES.

XX

XX Nelson EL, Nelson PJ;

XX

XX WPI; 1998-159552/14.

XX

XX Humanised polynucleotide vectors - comprising human derived promoter

XX and sequence acceptance site, used for the production of vaccines

XX

XX Example 4; Page 33-34; 125pp; English.

XX

XX Plasmid pITL-GFP comprises base vector pITL (see AAV21724) and a

XX humanised green fluorescent protein (GFP) reporter sequence

XX (see AAV21725). Novel humanised vectors comprise a human-derived

XX promoter or mammalian homologue which is functional in mammalian

XX target tissue and cells and a sequence acceptance site which

XX accepts cDNA products from RT-PCR cloning. pITL-GFP was used to

XX examine the kinetics of expression of a reporter sequence from

XX such vectors in an animal model (Fisher 344 rats). Toxicity from

CC polynucleotide vaccination was examined. The novel vectors are

CC used to express target antigens, especially tumour antigens. They

CC are non-replicating in mammalian cells but are capable of extended

CC stable expression of target sequences generating an immune response

CC in immunised individuals. The vectors selectively elicit immune

CC responses to the target sequences with little or no immune response

CC to the other components of the vectors. The target antigens are

CC expressed as intracellular polypeptides or peptides and, as such,

CC are processed as self polypeptides or peptides and appropriately

CC presented on antigen presenting cells.

XX

SQ Sequence 1911 BP; 474 A; 514 C; 511 G; 412 T; 0 other;

Query Match 67.0%; Score 152; DB 19; Length 1911;

Best Local Similarity 83.3%; Pred. NO. 1.9e-38;

Matches 210; Conservative 0; Mismatches 10; Indels 32; Gaps 2;

QY 8 GGGCCATATGGTGAGTGGATCCCTTGACCCCGAGCGGGGAT-GGGGAGACCTGAGTCAG 66

Db 716 gagccatatgtgagtgatgcttgaccccgagcgggatggggagacctgagtcag 775

QY 67 AGCCCCGGGCGACACAGCGCAATGCCCTCTCCCTCGCAG----- 109

Db 776 agccccgggcagcacagggccaatgccgtctctccctgcagtgagtagtgcgcg 835

QY 110 -----GATGAGTAGTAGTGCTCTCTCGGCCCTGGAAAGTGGCCACTCCAG 155

Db 836 ggtgggataccctgtgacccctccccagtgctctctctgcccgtgaagtggccactccag 895

QY 156 TGGCCACCAGCTTGCTCTAATAAATAAATTAAGTTGCACTATTGCTGACTAGTGCTCT 215

Db 896 tgccaccagcctgtctctataataataaattgagtgatcatttctgtgactgagtgctct 955

QY 216 CTATAATATTAT 227

Db 956 ctataatattat 967

RESULT 9

AAV21770

ID AAT76770 standard; cDNA; 2771 BP.

XX

AC AAT76770;

XX

DT 15-SBP-1997 (first entry)

XX

XX Rat Fabpl gut-specific promoter and human growth hormone exon 1.

XX

XX Human GDP-L-fucose:beta-D-galactoside-2-alpha-L-fucosyltransferase;

KW GDP-L-fucose:beta-D-N-acetylglucosamide 3/4-alpha-L-fucosyltransferase;

KW alpha 1,2 FT; alpha 1,3/4 FT; tissue-specific promoter;

KW rat liver fatty acid binding protein; transgene; transgenic mouse;

KW animal model; intestinal adhesion; Helicobacter pylori infection;

KW stomach; small intestine; gut; epithelial cell; surface receptor;

KW carbohydrate antigen; gastritis; peptic ulcer; neoplasia;

KW gastric adenocarcinoma; Lewis antigen; fucosylation; ds.

XX

XX Chimeric - Homo sapiens.

OS Chimeric - Rattus sp.

XX

XX Key Location/Qualifiers

FT promoter 1..617

FT FT /*tag= a

FT FT /note= "Corresponds to nucleotides -596 to +21 of

FT FT the rat liver fatty acid binding protein

FT FT gene"

FT FT 620..2771

FT FT /*tag= b

FT FT /number= 1

FT FT /note= "Corresponds to nucleotides +3 to +2150 of

FT FT the human growth hormone (hGH) gene; it is

FT FT desirable for cDNA coding for a human

FT fucosyltransferase to be inserted into
FT exon 1 of the hGH gene"

US5625124-A.

29-APR-1997.

11-JUL-1994; 94US-0273411.

11-JUL-1994; 94US-0273411.

(UNIW) UNIV WASHINGTON.

Falk P, Gordon JI;

WPI; 1997-258275/23.

Animal model for Helicobacter pylori infection - comprising
transgenic mouse expressing human enzyme promoting intestinal
adhesion

Example 2; Columns 25-28; 24pp; English.

A claimed transgenic mouse expresses, in its intestinal epithelial
cells, the enzyme human GDP-L-fucose:beta-D-galactoside 2-alpha-L-
fucosyltransferase (also called alpha 1,2 FT) or human GDP-L-fucose:
beta-D-N-acetylglucosamide 3/4-alpha-L-fucosyltransferase (also
called alpha 1,3/4 FT). The enzyme is expressed under the
control of a gut epithelial cell-specific promoter and Helicobacter
pylori adheres to the transgenic cells. The transgenic mouse and
intestinal epithelial cells from it are useful as models for screening
compounds for the ability to inhibit adhesion of H. pylori to gut
epithelial cells. The first 617 nucleotides of the present chimeric
sequence encode the promoter from rat liver fatty acid binding protein
(fabp1) which can direct foreign gene expression to the pit cell
lineage of the mouse gastric epithelium, to proliferating and non-
proliferating cells in intestinal crypts, as well as to the four
principal differentiated cell lineages along the crypt-to-villus axis
of the small intestine. The remainder of the chimeric sequence
corresponds to exon 1 of the human growth hormone (hGH) gene, into
which a sequence coding for a human fucosyltransferase (hFT) can be
inserted. No hGH will be produced because the initiator Met codon
and the first translation stop codon will be from the hFT sequence
and there is no ribosome re-entry sequence. The hGH exon 1 ensures
efficient splicing of the transgene primary transcript, improves
stability of the cytoplasmic hFT mRNA and allows transgene
expression to be monitored by in situ hybridisation using a
digoxigenin-labelled hGH oligonucleotide.

Sequence 2771 BP; 666 A; 718 C; 704 G; 683 T; 0 other;

Query Match 46.4%; Score 105.4; DB 18; Length 2771;
Best Local Similarity 99.1%; Pred. No. 1.3e-23;
Matches 106; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 121 AGTGCCTCTCTGGCCCTGGAGTTCGCACTCCAGTGCACACAGCCTTGCTCTAATAAA 180
|||||
DB 2177 agtgcctctcctggccttggaaagtggcaactccagtgccaccagccttgctcctaataaa 2236
|||||

QY 181 ATTAAGTTCATCATTTTCTGACTAGGTGCTCTATAAATTAT 227
|||||

DB 2237 attaatgttcacattttgtctgactaggtgctcctcctaataattat 2283
|||||

RESULT 10

AAC83268

ID AAC83268 standard; DNA; 4456 BP.

XX AAC83268;

AC AAC83268;

DT 16-MAR-2001 (first entry)

XX

DE Nucleotide sequence of HS2-beta globin-hGH transgene.

XX Erythropoietin-inducible transgene; human; sheep; juvenile beta globin;
KW human growth hormone; hGH; osteopathic; osteopaenia; osteoporosis;
KW hypersensitive element; HS; locus control region; ds.

OS Chimeric - Homo sapiens.

OS Chimeric - Ovis aries.

PN US6153427-A.

XX 28-NOV-2000.

XX 12-OCT-1994; 94US-0321686.

XX 12-OCT-1994; 94US-0321686.

PA (UYNE-) UNIV NORTHEASTERN OHIO.

XX Schneider GB, King D;

XX WPI; 2001-060092/07.

XX New DNA construct comprising an erythropoietin-inducible sequence, an
PT erythropoietin-inducible sequence and coding sequence of interest
PT useful for treating osteopenia, particularly osteoporosis -

PS Claim 10; Fig 1B; 19pp; English.

XX This invention relates to a DNA construct comprising an

CC erythropoietin-inducible sequence represented by the present sequence. The
CC HS2-beta globin-hGH transgene comprises the hypersensitive site (HS2)
CC element from the human beta-globin locus control region, the juvenile
CC beta globin promoter isolated from sheep, and the human growth hormone
CC (hGH) gene. The transgene results in osteopathic activity, when used in
CC gene therapy. The DNA construct is useful for the treatment of
CC osteopaenias, such as osteoporosis.

XX Sequence 4456 BP; 1238 A; 978 C; 1076 G; 1164 T; 0 other;

Query Match 46.4%; Score 105.4; DB 22; Length 4456;
Best Local Similarity 99.1%; Pred. No. 1.6e-23;
Matches 106; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 121 AGTGCCTCTCTGGCCCTGGAGTTCGCACTCCAGTGCACACAGCCTTGCTCTAATAAA 180
|||||

DB 4324 agtgcctctcctggccttggaaagtggcaactccagtgccaccagccttgctcctaataaa 4383
|||||

QY 181 ATTAAGTTCATCATTTTCTGACTAGGTGCTCTATAAATTAT 227
|||||

DB 4384 attaatgttcacattttgtctgactaggtgctcctcctaataattat 4430
|||||

RESULT 11

AAV33343

ID AAV33343 standard; DNA; 191 BP.

XX AAV33343;

XX 18-NOV-1998 (first entry)

XX Human growth hormone 3' UTR/poly(A) signal sequence.

XX Human interleukin 2; IL-2; CMV promoter; cationic lipid; DOTMA;
KW human growth hormone 3'-untranslated region; lipid cholesterol;
KW gene therapy; tumour; helper T cell; fever; fluid retention;
KW vascular leak syndrome; human growth hormone 3' UTR; ss.

OS Homo sapiens.

XX Key

FH poly_a_signal

FT Location/Qualifiers

81..86

FT XX /*Lag= a
 PN WO9834952-A2.
 PD 13-AUG-1998.
 XX 09-FEB-1998; 98WO-US02221.
 XX 23-JAN-1998; 98US-0012366.
 PR 10-FEB-1997; 97US-0039709.
 XX (GENE-) GENEMEDICINE INC.
 PA Bruno M, Muller S, Mumper R, Munger W, Ralston R;
 PI WPI; 1998-467159/40.
 DR Plasmid for expression of a human IL-2 coding sequence - and
 XX lipid/DNA delivery systems which have suitable positive/negative
 PT charge ratios and are useful in treating tumours
 PT
 XX Disclosure; Page 23; 77pp; English.
 PS
 XX The present sequence represents a human growth hormone 3' UTR/poly(A)
 CC signal sequence. The invention provides a pIL0697 plasmid comprising
 CC a CMV promoter/enhancer transcriptionally linked to a codon
 CC optimised human IL-2 coding sequence (AAV3342) and the present human
 CC growth hormone 3'-untranslated region/poly(A) signal. Also, a 5' UTR
 CC (AAV3344) may be present between the 3' of the promoter region and the
 CC 5' of the IL-2 coding region. The invention also provides a composition
 CC for delivery and expression of a human IL-2 coding sequence in mammals,
 CC comprising: (a) a cationic lipid e.g. DOTMA, a neutral co-lipid e.g.
 CC lipid cholesterol and (b) a plasmid as described above. The
 CC plasmid/composition may be used in gene therapy, especially therapy of
 CC tumours, as IL-2 is involved in stimulating proliferation of helper T
 CC cells. The claimed antitumour effect of the above formulations does
 CC not only depend on the expression level of IL-2 but is also a function
 CC of the non-DNA formulation components, with the effect being greater
 CC than merely additive. Also, as administration of high doses of IL-2
 CC protein results in significant toxicity-related side effects, e.g.
 CC fever, fluid retention and vascular leak syndrome, the present
 CC composition is claimed to avoid such problems.
 XX
 SQ Sequence 191 BP; 39 A; 47 C; 56 G; 49 T; 0 other;

Query Match 42.3%; Score 96; DB 19; Length 191;
 Best Local Similarity 99.1%; Pred. No. 4.5e-21;
 Matches 107; Conservative 0; Mismatches 1; Gaps 1;

OY 121 AGTGCCTCTCTGCGCCCTGGAGTTGCCACTCCAGTGCACGCCAGCCTGTCTTAATAAA 180
 |||||
 DB 27 agtgcctctctgcccctgggaagttgccactccagtgcccaccagcctgtcttaataaa 86
 |||||

OY 181 ATTAAGTTGCATCATTTTGTCTGCTAGTGTGCC-TCTATAATATAT 227
 |||||
 DB 87 attaatgtgcatactttgtctgactaggtgtcctctataataattat 134
 |||||

RESULT 12
 AAZ40410
 ID AAZ40410 standard; DNA; 191 BP.
 XX
 AC AAZ40410;
 XX
 DT 15-FEB-2000 (first entry)
 DE Human growth hormone 3' UTR sequence.
 XX
 KW Wild type; human; interferon-alpha; plasmid; cytomegalovirus; CMV;
 KW promoter; growth hormone; untranslated region; UTR; mammal; disease;
 KW cancer; ss.

OS Homo sapiens.
 XX
 PN WO9947678-A2.
 XX
 PD 23-SEP-1999.
 XX
 XX 12-MAR-1999; 99WO-US05394.
 XX
 XX 19-MAR-1998; 98US-0078654.
 XX
 XX (GENE-) GENEMEDICINE INC.
 PA Nordstrom J, Pericle F, Rolland A, Ralston R;
 PI WPI; 1999-562116/47.
 DR
 XX New plasmids containing an interferon-alpha coding sequence, used for
 PT the treatment of a mammalian condition of disease, particularly cancer
 PT
 XX
 XX Disclosure; Page 27; 137pp; English.
 PS
 XX This sequence represents the 3' untranslated region from a human
 CC growth hormone gene. The invention relates to a novel plasmid comprising
 CC a cytomegalovirus (CMV) promoter transcriptionally linked with an
 CC interferon alpha (IFN-alpha) coding sequence, and a growth hormone
 CC 3'-untranslated region (UTR). The plasmids can be used for treating
 CC a mammalian condition of disease, particularly cancer.
 XX
 SQ Sequence 191 BP; 39 A; 47 C; 56 G; 49 T; 0 other;

Query Match 42.3%; Score 96; DB 20; Length 191;
 Best Local Similarity 99.1%; Pred. No. 4.5e-21;
 Matches 107; Conservative 0; Mismatches 0; Indels 1; Gaps 1;

OY 121 AGTGCCTCTCTGCGCCCTGGAGTTGCCACTCCAGTGCACGCCAGCCTGTCTTAATAAA 180
 |||||
 DB 27 agtgcctctctgcccctgggaagttgccactccagtgcccaccagcctgtcttaataaa 86
 |||||

OY 181 ATTAAGTTGCATCATTTTGTCTGCTAGTGTGCC-TCTATAATATAT 227
 |||||
 DB 87 attaatgtgcatactttgtctgactaggtgtcctctataataattat 134
 |||||

RESULT 13
 AAZ50393
 ID AAZ50393 standard; DNA; 191 BP.
 XX
 AC AAZ50393;
 XX
 DT 18-MAY-2000 (first entry)
 DE Human growth hormone 3' UTR.
 XX
 XX Human growth hormone 3' UTR; anti-angiogenic agent; expression plasmid;
 KW cancer; gene expression; mRNA stability; transfection; tumour activity;
 KW cytosstatic; gene therapy; lung metastatic tumour; ss.
 XX
 OS Homo sapiens.
 XX
 FH Key Location/Qualifiers
 XX 3' UTR 1..100
 FT /*tag= a
 FT polyA_signal 81..86
 FT /*tag= b
 XX
 XX WO200006759-A2.
 XX
 PD 10-FEB-2000.
 XX
 XX 20-JUL-1999; 99WO-US16388.

PR 27-JUL-1998; 98US-0094375.
 XX (VALE-) VALENTIS INC.
 XX
 PI Min W, Szymanski P, Mehrens D, Ralston R, Sullivan S;
 XX WPI; 2000-183133/16.
 XX
 PT Plasmids comprising tissue specific transcription elements linked to an
 PT anti-angiogenic gene is useful transfection of cells and treatment of,
 PT e.g. Cancer -
 XX
 PS Disclosure; Page 28; 103pp; English.
 XX
 CC The present sequence is the 3'UTR of human growth hormone along with
 CC 3'flanking sequence. This sequence can be linked immediately following
 CC natural translation termination codon for a nucleotide encoding anti
 CC -angiogenic agent. This was used in the construction of expression
 CC plasmid incorporating an anti-angiogenic agent for the treatment of
 CC mammalian diseases, especially cancer. This 3'UTR influences gene
 CC expression by controlling the accuracy and efficiency of RNA processing,
 CC mRNA stability and translation. The plasmids can be used for (in vivo)
 CC transfection of a cell in situ in order to modulate tumour activity.
 CC Anti-angiogenic gene inhibits growth of solid tumour and lung metastatic
 CC tumours by intravenous or intramuscular delivery.
 XX
 SQ Sequence 191 BP; 39 A; 47 C; 56 G; 49 T; 0 other;

Query Match 42.3%; Score 96; DB 21; Length 191;
 Best Local Similarity 99.1%; Pred. No. 4.5e-21;
 Matches 107; Conservative 0; Mismatches 0; Indels 1; Gaps 1;
 Qy 121 AGTGGCTCTCTGCGCCCTGGAAGTTGCCACTTCACAGTCCGCCACCGCCCTTGTCTTAATAAA 180
 Db 27 agtgcctctctgcccctggaagtggcactccagtgccaccagcctgtcctaataaa 86
 Qy 181 ATTAAGTTGCATCATTTTGTCTGACTAGTGTGCC-TCTATATATTAT 227
 Db 87 attaatgtgcacattttgtctgactaggtgtcctctcctaataattat 134

RESULT 14
 AAF76860
 ID AAF76860 standard; cDNA; 563 BP.
 XX
 AC AAF76860;
 XX
 DT 14-MAY-2001 (first entry)
 XX
 DE Human secreted protein cDNA #18.
 XX
 KW Human; secreted protein; immunomodulatory; antisclerotic;
 KW dermatological; immunosuppressive; antiinflammatory; anti-HIV;
 KW immunostimulant; cytostatic; cardiac; vascular; anti-angiogenic;
 KW ophthalmological; neuroprotectant; nootropic; anticonvulsant; vaccine;
 KW antialzheimers; antiparkinsonian; antimicrobial; vulnery; gene therapy;
 KW immune disorder; hyperproliferative; cardiovascular; angiogenic;
 KW neurological; infection; ss.
 XX
 OS Homo sapiens.
 XX
 PN WO200112776-A2.
 XX
 PD 22-FEB-2001.
 XX
 PF 15-AUG-2000; 2000WO-US22350.
 XX
 PR 16-AUG-1999; 99US-0148759..
 XX
 PA (HUMA-) HUMAN GENOME SCI INC.
 XX
 PI Shi Y, Young PE, Ebner R, Soppet DR, Ruben SM;

XX WPI; 2001-244245/25.
 DR P-PSDB; AAB70079.
 XX
 PT Nucleic acids encoding 18 human secreted polypeptides, useful for
 PT preventing, diagnosing and/or treating e.g. cancers, Parkinson's
 PT disease and diabetic retinopathy -
 XX
 PS Claim 1; Page 358; 380pp; English.
 XX
 CC The present sequence is one of 18 nucleic acid molecules encoding novel
 CC human secreted proteins. The nucleic acids and proteins may be used in
 CC the prevention, diagnosis and treatment of diseases including immune
 CC disorders (e.g. multiple sclerosis, systemic lupus erythematosus and
 CC human immunodeficiency virus (HIV) infections), hyperproliferative
 CC disorders (e.g. cancers and Gaucher's disease), cardiovascular diseases
 CC (e.g. Schmitz syndrome, Chaga's cardiomyopathy and coronary
 CC arteriosclerosis), angiogenic disorders (e.g. corneal graft
 CC neovascularisation and diabetic retinopathy), neurological disorders
 CC (e.g. Huntington's chorea, Alzheimer's disease and Parkinson's disease),
 CC infectious diseases and/or for promoting wound healing, regeneration
 CC and/or chemotaxis. The nucleic acid molecules may also be used as DNA
 CC probes in diagnostic assays to detect and quantitate the presence of
 CC similar nucleic acid sequences in samples. The polypeptides may also be
 CC used as antigens in the production of antibodies and in assays to
 CC identify modulators of protein expression and activity.
 XX
 SQ Sequence 563 BP; 127 A; 139 C; 176 G; 121 T; 0 other;

Query Match 42.3%; Score 96; DB 22; Length 563;
 Best Local Similarity 99.1%; Pred. No. 6.8e-21;
 Matches 107; Conservative 0; Mismatches 0; Indels 1; Gaps 1;
 Qy 121 AGTGGCTCTCTGCGCCCTGGAAGTTGCCACTTCACAGTCCGCCACCGCCCTTGTCTTAATAAA 180
 Db 349 agtgcctctctgcccctggaagtggcactccagtgccaccagcctgtcctaataaa 408
 Qy 181 ATTAAGTTGCATCATTTTGTCTGACTAGTGTGCC-TCTATATATTAT 227
 Db 409 attaatgtgcacattttgtctgactaggtgtcctctcctaataattat 456

RESULT 15
 AAX15886
 ID AAX15886 standard; DNA; 1441 BP.
 XX
 AC AAX15886;
 XX
 DT 13-MAY-1999 (first entry)
 XX
 DE NotI/SalI psk-GHRH plasmid fragment containing GHRH DNA sequence.
 XX
 KW Human; growth hormone releasing hormone; GHRH; osteoporosis; cachexia;
 KW growth disorder; burn; sepsis; trauma; atherogenic; cardiovascular;
 KW chronic obstructive pulmonary disease; atherosclerotic; cerebrovascular;
 KW peripheral vascular disease; haemophilia; muscular atrophy; aging;
 KW muscular dystrophy; ss.
 XX
 OS Synthetic.
 OS Homo sapiens.
 XX
 PN WO9905300-A2.
 XX
 PD 04-FEB-1999.
 XX
 PF 24-JUL-1998; 98WO-US15434.
 XX
 PR 20-OCT-1997; 97US-0062608.
 PR 24-JUL-1997; 97US-0053609.
 XX
 PA (GENE-) GENEMEDICINE INC.
 XX

Search completed: January 17, 2002, 12:02:18
Job time: 17725 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: January 17, 2002, 07:04:33 ; Search time 221.34 Seconds
(without alignments)
232.269 Million cell updates/sec

Title: US-09-242-202A-10
Perfect score: 227
Sequence: 1 GCCTTAAGGCCATATGGTG.....GGTGCTCTATAATATAT 227

Scoring table: IDENTITY NUC
Gapop 10.0 , Gapext 1.0

Searched: 351203 seqs, 113238999 residues
Total number of hits satisfying chosen parameters: 702406

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued_Patents_NA.*
1: /cgn2_6/ptodata/2/ina/5A_COMB.seq.*
2: /cgn2_6/ptodata/2/ina/5B_COMB.seq.*
3: /cgn2_6/ptodata/2/ina/6A_COMB.seq.*
4: /cgn2_6/ptodata/2/ina/6B_COMB.seq.*
5: /cgn2_6/ptodata/2/ina/PTUS_COMB.seq.*
6: /cgn2_6/ptodata/2/ina/backfiles1.seq.*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	105.4	46.4	2771	1 US-08-273-411-5	Sequence 5, Appli
2	105.4	46.4	4456	3 US-08-321-686B-1	Sequence 1, Appli
3	96	42.3	191	3 US-09-012-366-4	Sequence 4, Appli
4	96	42.3	2086	3 US-08-589-028-9	Sequence 9, Appli
5	96	42.3	2086	3 US-08-784-582-9	Sequence 9, Appli
6	96	42.3	2086	4 US-08-785-271-9	Sequence 9, Appli
7	93.4	41.1	3653	4 US-08-973-334-1	Sequence 1, Appli
8	93.4	41.1	3653	4 US-09-563-869A-1	Sequence 1, Appli
9	93.4	41.1	3653	4 US-08-549-489-1	Sequence 1, Appli
10	72	31.7	792	3 US-08-887-674E-9	Sequence 9, Appli
11	60.6	26.7	1884	3 US-08-784-582-70	Sequence 70, Appl
12	60.6	26.7	2356	3 US-08-784-582-72	Sequence 72, Appl
c 13	42	18.5	42	3 US-08-068-754-3	Sequence 3, Appli
c 14	42	18.5	42	4 US-09-325-926-3	Sequence 3, Appli
c 15	42	18.5	51	2 US-08-453-024-10	Sequence 10, Appl
c 16	42	18.5	51	4 US-08-070-162-10	Sequence 10, Appl
c 17	39	17.2	46	4 US-09-173-043-10	Sequence 10, Appl
c 18	39	17.2	46	4 US-09-209-525-15	Sequence 15, Appl
19	39	17.2	153	4 US-09-209-525-16	Sequence 16, Appl
20	32.6	14.4	755	1 US-08-468-824-5	Sequence 5, Appli
c 21	31.8	14.0	3331	4 US-09-042-785A-1	Sequence 1, Appli
22	31.2	13.7	1758	3 US-09-191-171-6	Sequence 6, Appli
23	31.2	13.7	1758	4 US-09-385-707-6	Sequence 6, Appli
24	31.2	13.7	2220	2 US-08-864-224-1	Sequence 1, Appli
25	31.2	13.7	3853	3 US-08-801-092-5	Sequence 5, Appli
26	31.2	13.7	4026	3 US-08-801-092-19	Sequence 19, Appl
27	31.2	13.7	4249	3 US-08-801-092-33	Sequence 33, Appl

28	31.2	13.7	4283	1 US-08-343-401A-3	Sequence 3, Appli
29	31.2	13.7	4283	1 US-08-445-265A-1	Sequence 1, Appli
30	31.2	13.7	4283	3 US-08-990-442-1	Sequence 1, Appli
31	31.2	13.7	4326	4 US-08-760-615-7	Sequence 7, Appli
32	31.2	13.7	4328	4 US-09-132-808-1	Sequence 1, Appli
33	31.2	13.7	4328	4 US-08-910-647-2	Sequence 2, Appli
34	31.2	13.7	4328	4 US-08-910-647-2	Sequence 4, Appli
35	31.2	13.7	4928	1 US-08-345-913-1	Sequence 1, Appli
36	31.2	13.7	4928	3 US-08-818-562-1	Sequence 1, Appli
37	31.2	13.7	4965	2 US-08-564-313-1	Sequence 1, Appli
38	31.2	13.7	4965	5 PCT-US94-06089-1	Sequence 1, Appli
39	31.2	13.7	5107	4 US-08-910-647-3	Sequence 3, Appli
40	31.2	13.7	5653	1 US-08-073-836-3	Sequence 3, Appli
41	31.2	13.7	5653	1 US-08-235-277-1	Sequence 1, Appli
42	31.2	13.7	5676	2 US-08-663-998-3	Sequence 3, Appli
43	31.2	13.7	5682	2 US-08-663-998-4	Sequence 4, Appli
44	31.2	13.7	5703	1 US-08-467-420A-50	Sequence 50, Appl
45	31.2	13.7	5703	1 US-08-470-110A-50	Sequence 50, Appl

ALIGNMENTS

RESULT 1
US-08-273-411-5
; Sequence 5, Application US/08273411
; Patent No. 5625124
; GENERAL INFORMATION:
; APPLICANT: Falk, Per
; APPLICANT: Gordon, Jeffrey I.
; TITLE OF INVENTION: Animal Model for Gastro-Intestinal
; TITLE OF INVENTION: Disease
; NUMBER OF SEQUENCES: 10
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Patrea L. Pabst
; STREET: 1100 Peachtree Street, Suite 2800
; CITY: Atlanta
; STATE: Georgia
; COUNTRY: USA
; ZIP: 30309-4530
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/273,411
; FILING DATE:
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Pabst, Patrea L.
; REGISTRATION NUMBER: 31,284
; REFERENCE/DOCKET NUMBER: WU106
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (404) 815-6508
; TELEFAX: (404) 815-6555
; INFORMATION FOR SEQ ID NO: 5:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2771 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
; PUBLICATION INFORMATION:
; AUTHORS: Sweetser, et al.
; JOURNAL: Genes & Dev.
; VOLUME: 2
; PAGES: 1318-1332
; DATE: 1988
; RELEVANT RESIDUES IN SEQ ID NO: 5: FROM 561 TO 629
; PUBLICATION INFORMATION:

```

; AUTHORS: Seeburg, et al.
; JOURNAL: DNA
; VOLUME: 1
; PAGES: 239-249
; DATE: 1982
; RELEVANT RESIDUES IN SEQ ID NO: 5: FROM 620 TO 2771
; PUBLICATION INFORMATION:
; AUTHORS: Sweetser, et al.
; JOURNAL: J. Biol. Chem.
; VOLUME: 261
; PAGES: 5553-5561
; DATE: 1986
; RELEVANT RESIDUES IN SEQ ID NO: 5: FROM 1 TO 617
; US-08-273-411-5
;
; Query Match 46.4%; Score 105.4; DB 1; Length 2771;
; Best Local Similarity 99.1%; Pred. No. 5.4e-24;
; Matches 106; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
;
QY 121 AGTGCCTCTCTGGCCCTGGAAGTTGCCACTCCAGTGGCCACGAGCCTTGCTCTAATAAAA 180
Db 2177 AGTGCCTCTCTGGCCCTGGAAGTTGCCACTCCAGTGGCCACGAGCCTTGCTCTAATAAAA 2236
;
QY 181 ATTAAGTTGCATCATTTGCTGACTAGTGTCTCTCTAATAATATAT 227
Db 2237 ATTAAGTTGCATCATTTGCTGACTAGTGTCTCTCTAATAATATAT 2283
;
RESULT 2
US-08-321-686B-1
; Sequence 1, Application US/08321686B
; Patent No. 6153427
; GENERAL INFORMATION:
; APPLICANT: King, Donna
; APPLICANT: Schneider, Gary B.
; TITLE OF INVENTION: Treatment of
; TITLE OF INVENTION: Osteopenias and No. 6153427-Human Transgenic
; TITLE OF INVENTION: Mammals Expressing Therapeutic
; NUMBER OF SEQUENCES: 1
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Wallenstein & Wagner, Ltd.
; STREET: 311 South Wacker Drive
; CITY: Chicago
; STATE: Illinois
; COUNTRY: USA
; ZIP: 60606
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 3.50 inch
; MEDIUM TYPE: 800 Kb storage
; COMPUTER: IBM compatible
; OPERATING SYSTEM: WordPerfect 5.1
; SOFTWARE: Simple Text ASCII (IBM format)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/321,686B
; FILING DATE: October 12, 1994
; CLASSIFICATION: 800
; PRIOR APPLICATION DATA: not
; PRIOR APPLICATION DATA: applicable
; ATTORNEY/AGENT INFORMATION:
; NAME: Barry, Alan L.
; REGISTRATION NUMBER: 30,819
; REFERENCE/DOCKET NUMBER: 1017P021
; TELEPHONE: (312)554-3300
; TELEFAX: (312)554-3301
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 4456 bp
; TYPE: nucleic acid
; STRANDEDNESS: single

```

```

; TOPOLOGY: linear
; US-08-321-686B-1
;
; Query Match 46.4%; Score 105.4; DB 3; Length 4456;
; Best Local Similarity 99.1%; Pred. No. 6.4e-24;
; Matches 106; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
;
QY 121 AGTGCCTCTCTGGCCCTGGAAGTTGCCACTCCAGTGGCCACGAGCCTTGCTCTAATAAAA 180
Db 4324 AGTGCCTCTCTGGCCCTGGAAGTTGCCACTCCAGTGGCCACGAGCCTTGCTCTAATAAAA 4383
;
QY 181 ATTAAGTTGCATCATTTGCTGACTAGTGTCTCTCTAATAATATAT 227
Db 4384 ATTAAGTTGCATCATTTGCTGACTAGTGTCTCTCTAATAATATAT 4430
;
RESULT 3
US-09-012-366-4
; Sequence 4, Application US/09012366
; Patent No. 6034072
; GENERAL INFORMATION:
; APPLICANT: Robert Ralston
; APPLICANT: Susanne Muller
; APPLICANT: Russ Mumper
; APPLICANT: William Munger
; APPLICANT: Maria Bruno
; TITLE OF INVENTION: IL-2 GENE EXPRESSION AND
; TITLE OF INVENTION: DELIVERY SYSTEMS AND USES
; NUMBER OF SEQUENCES: 11
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Lyon & Lyon
; STREET: 633 West Fifth Street
; CITY: Los Angeles
; STATE: California
; COUNTRY: U.S.A.
; ZIP: 90071-2066
; COMPUTER READABLE FORM:
; MEDIUM TYPE: 3.5" Diskette, 1.44 Mb
; MEDIUM TYPE: storage
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: IBM P.C. DOS 5.0
; SOFTWARE: FastSeq for Windows 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/012,366
; FILING DATE: January 23, 1998
; CLASSIFICATION: 424
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 60/039,709
; FILING DATE: February 10, 1997
; ATTORNEY/AGENT INFORMATION:
; NAME: Berkman, Charles S.
; REGISTRATION NUMBER: 38,077
; REFERENCE/DOCKET NUMBER: 230/214
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (213) 489-1600
; TELEFAX: (213) 955-0440
; TELEX: 67-3510
; INFORMATION FOR SEQ ID NO: 4:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 191 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; US-09-012-366-4
;
; Query Match 42.3%; Score 96; DB 3; Length 191;
; Best Local Similarity 99.1%; Pred. No. 1.8e-21;
; Matches 107; Conservative 0; Mismatches 0; Indels 1; Gaps 1;
;
QY 121 AGTGCCTCTCTGGCCCTGGAAGTTGCCACTCCAGTGGCCACGAGCCTTGCTCTAATAAAA 180

```


Db 27 AGTGCCTCTCCTGGCCCTGGAAGTTGCCACTCCAGTGGCCACCAGCCTTGCTCTAATAAA 86
 QY 181 ATTAAGTTCATCATTTCTCTGACTAGGTGCC-TCTATAATATAT 227
 Db 87 ATTAAGTTCATCATTTCTCTGACTAGGTGCC-TCTATAATATAT 134

RESULT 4
 US-08-589-028-9
 ; Sequence 9, Application US/08589028
 ; Patent No. 6087129
 ; GENERAL INFORMATION:
 ; APPLICANT: Newgard, Christopher B.
 ; APPLICANT: Halban, Philippe A.
 ; APPLICANT: No. 6087129mington, Karl D.
 ; APPLICANT: Clark, Samuel A.
 ; APPLICANT: Thigpen, Anice E.
 ; APPLICANT: Quaade, Christian
 ; APPLICANT: Kruse, Fred
 ; TITLE OF INVENTION: Recombinant Expression of Proteins From
 ; NUMBER OF SEQUENCES: 50
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: Arnold, White & Durkee
 ; STREET: P. O. Box 4433
 ; CITY: Houston
 ; STATE: TX
 ; COUNTRY: USA
 ; ZIP: 77210-4433
 ; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: Floppy disk
 ; COMPUTER: IBM PC compatible
 ; OPERATING SYSTEM: PC-DOS/MS-DOS
 ; SOFTWARE: PatentIn Release #1.0, Version #1.30
 ; CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: US/08/589,028
 ; FILING DATE: Concurrently Herewith
 ; CLASSIFICATION: 435
 ; ATTORNEY/AGENT INFORMATION:
 ; NAME: Highlander, Steven L.
 ; REGISTRATION NUMBER: 47,642
 ; REFERENCE/DOCKET NUMBER: UTSD:426\HYL
 ; TELECOMMUNICATION INFORMATION:
 ; TELEPHONE: (512) 418-3000
 ; TELEFAX: (512) 474-7577
 ; INFORMATION FOR SEQ ID NO: 9:
 ; SEQUENCE CHARACTERISTICS:
 ; LENGTH: 2086 base pairs
 ; TYPE: nucleic acid
 ; STRANDEDNESS: single
 ; TOPOLOGY: linear

US-08-589-028-9
 Query Match 42.3%; Score 96; DB 3; Length 2086;
 Best Local Similarity 99.1%; Pred No. 4.4e-21;
 Matches 107; Conservative 0; Mismatches 0; Indels 1; Gaps 1;
 QY 121 AGTGCCTCTCCTGGCCCTGGAAGTTGCCACTCCAGTGGCCACCAGCCTTGCTCTAATAAA 180
 Db 1559 AGTGCCTCTCCTGGCCCTGGAAGTTGCCACTCCAGTGGCCACCAGCCTTGCTCTAATAAA 1618
 QY 181 ATTAAGTTCATCATTTCTCTGACTAGGTGCC-TCTATAATATAT 227
 Db 1619 ATTAAGTTCATCATTTCTCTGACTAGGTGCC-TCTATAATATAT 1666

RESULT 5
 US-08-784-582-9
 ; Sequence 9, Application US/08784582
 ; Patent No. 6110707
 ; GENERAL INFORMATION:
 ; APPLICANT: Newgard, Christopher B.
 ; APPLICANT: Halban, Philippe A.
 ; APPLICANT: No. 6194176mington, Karl D.
 ; APPLICANT: Clark, Samuel A.
 ; APPLICANT: Thigpen, Anice E.
 ; APPLICANT: Quaade, Christian

APPLICANT: Newgard, Christopher B.
 APPLICANT: Halban, Philippe A.
 APPLICANT: No. 6110707mington, Karl D.
 APPLICANT: Clark, Samuel A.
 APPLICANT: Thigpen, Anice E.
 APPLICANT: Quaade, Christian
 APPLICANT: Kruse, Fred
 APPLICANT: McGarry, Dennis
 TITLE OF INVENTION: RECOMBINANT EXPRESSION OF PROTEINS FROM
 NUMBER OF SEQUENCES: 79
 CORRESPONDENCE ADDRESS:
 ADDRESSEE: Arnold, White & Durkee
 STREET: P.O. Box 4433
 CITY: Houston
 STATE: Texas
 COUNTRY: USA
 ZIP: 77210
 COMPUTER READABLE FORM:
 MEDIUM TYPE: Floppy disk
 COMPUTER: IBM PC compatible
 OPERATING SYSTEM: PC-DOS/MS-DOS
 SOFTWARE: PatentIn Release #1.0, Version #1.30
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/08/784,582
 FILING DATE: Concurrently Herewith
 CLASSIFICATION: 435
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: US 60/028,427
 FILING DATE: 15-OCT-1996
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: US 08/589,028
 FILING DATE: 19-JAN-1996
 ATTORNEY/AGENT INFORMATION:
 NAME: Highlander, Steven L.
 REGISTRATION NUMBER: 37,642
 REFERENCE/DOCKET NUMBER: UTSD:514
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: 512/418-3000
 TELEFAX: 512/474-7577
 INFORMATION FOR SEQ ID NO: 9:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 2086 base pairs
 TYPE: nucleic acid
 STRANDEDNESS: single
 TOPOLOGY: linear
 US-08-784-582-9

Query Match 42.3%; Score 96; DB 3; Length 2086;
 Best Local Similarity 99.1%; Pred No. 4.4e-21;
 Matches 107; Conservative 0; Mismatches 0; Indels 1; Gaps 1;
 QY 121 AGTGCCTCTCCTGGCCCTGGAAGTTGCCACTCCAGTGGCCACCAGCCTTGCTCTAATAAA 180
 Db 1559 AGTGCCTCTCCTGGCCCTGGAAGTTGCCACTCCAGTGGCCACCAGCCTTGCTCTAATAAA 1618
 QY 181 ATTAAGTTCATCATTTCTCTGACTAGGTGCC-TCTATAATATAT 227
 Db 1619 ATTAAGTTCATCATTTCTCTGACTAGGTGCC-TCTATAATATAT 1666

RESULT 6
 US-08-785-271-9
 ; Sequence 9, Application US/08785271
 ; Patent No. 6194176
 ; GENERAL INFORMATION:
 ; APPLICANT: Newgard, Christopher B.
 ; APPLICANT: Halban, Philippe A.
 ; APPLICANT: No. 6194176mington, Karl D.
 ; APPLICANT: Clark, Samuel A.
 ; APPLICANT: Thigpen, Anice E.
 ; APPLICANT: Quaade, Christian

; APPLICATION NUMBER: US/09/563,869A
; FILING DATE: 03-May-2000
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/973,334
; FILING DATE: <Unknown>
; APPLICATION NUMBER: US 08/549,489
; FILING DATE: 27-OCT-1995
; ATTORNEY/AGENT INFORMATION:
; NAME: Bak, Mary E.
; REGISTRATION NUMBER: 31,215
; REFERENCE/DOCKET NUMBER: GNPVN012CIPUSA
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (215) 540-9206
; TELEFAX: (215) 540-5818
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 3653 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: not relevant
; MOLECULE TYPE: cDNA
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 1521..2405
; SEQUENCE DESCRIPTION: SEQ ID NO: 1:
US-09-563-869A-1

Query Match 41.1%; Score 93.4; DB 4; Length 3653;
Best Local Similarity 97.2%; Pred. No. 3.6e-20;
Matches 105; Conservative 0; Mismatches 2; Indels 1; Gaps 1;

OY 121 AGTGCTCTCTCGGCCCTGGAAGTTGCCACTCCAGTGCCACCGCCAGCCCTGTCTCTAATAAA 180
|||||
DB 2501 AGTGCTCTCTCGGCCCTGGAAGTTGCCACTCCAGTGCCACCGCCAGCCCTGTCTCTAATAAA 2560
|||||

OY 181 ATTAAGTTGCATCATTTTCTCTGACTAGTGTCCTCTCTAATAATATAT 227
|||||
DB 2561 ATTAAGTTGNATCATTTTCTCTGACTAGTGTCCTCTCTAATAATATAT 2608
|||||

RESULT 9
US-08-549-489-1
; Sequence 1, Application US/08549489
; Patent No. 6281010
; GENERAL INFORMATION:
; APPLICANT: Wilson, James M.
; APPLICANT: Gao, Guang-Ping
; TITLE OF INVENTION: No. 6281010el Adenovirus Gene Therapy Vehicle
; TITLE OF INVENTION: and Cell Line
; NUMBER OF SEQUENCES: 3
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Howson and Howson
; STREET: Box 457, 321 No. 6281010ristown Road
; CITY: Spring House
; STATE: PA
; COUNTRY: USA
; ZIP: 19477
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/549,489
; FILING DATE:
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/462,014
; FILING DATE: 08-JUN-1995
; ATTORNEY/AGENT INFORMATION:
; NAME: Bak, Mary E.

; REGISTRATION NUMBER: 31,215
; REFERENCE/DOCKET NUMBER: GNPVN013
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (215) 540-9206
; TELEFAX: (215) 540-5818
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 3653 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: not relevant
; MOLECULE TYPE: cDNA
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 1521..2405
; SEQUENCE DESCRIPTION: SEQ ID NO: 1:
US-08-549-489-1

Query Match 41.1%; Score 93.4; DB 4; Length 3653;
Best Local Similarity 97.2%; Pred. No. 3.6e-20;
Matches 105; Conservative 0; Mismatches 2; Indels 1; Gaps 1;

OY 121 AGTGCTCTCTCGGCCCTGGAAGTTGCCACTCCAGTGCCACCGCCAGCCCTGTCTCTAATAAA 180
|||||
DB 2501 AGTGCTCTCTCGGCCCTGGAAGTTGCCACTCCAGTGCCACCGCCAGCCCTGTCTCTAATAAA 2560
|||||

OY 181 ATTAAGTTGCATCATTTTCTCTGACTAGTGTCCTCTCTAATAATATAT 227
|||||
DB 2561 ATTAAGTTGNATCATTTTCTCTGACTAGTGTCCTCTCTAATAATATAT 2608
|||||

RESULT 10
US-08-887-674E-9
; Sequence 9, Application US/08887674E
; Patent No. 6130092
; GENERAL INFORMATION:
; APPLICANT: Andre Lieber, Michael Strauss
; TITLE OF INVENTION: RIBOZYME GENE LIBRARY AND PRO-
; TITLE OF INVENTION: CESS FOR PREPARATION
; NUMBER OF SEQUENCES: 12
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Gabriel P. Katona
; STREET: 230 Park Avenue, Suite 2200
; CITY: New York
; STATE: NY
; COUNTRY: United States of America
; ZIP: 10169
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25 (EPO)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/887,674E
; FILING DATE: July 3, 1997
; INFORMATION FOR SEQ ID NO: 9:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 792 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: tRNA
US-08-887-674E-9

Query Match 31.7%; Score 72; DB 3; Length 792;
Best Local Similarity 100.0%; Pred. No. 1.1e-13;
Matches 72; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 121 AGTGCTCTCTCGGCCCTGGAAGTTGCCACTCCAGTGCCACCGCCAGCCCTGTCTCTAATAAA 180
|||||
DB 721 AGTGCTCTCTCGGCCCTGGAAGTTGCCACTCCAGTGCCACCGCCAGCCCTGTCTCTAATAAA 780
|||||

Qy 181 ATTAAGTTGCAT 192
|||||
Db 781 ATTAAGTTGCAT 792

RESULT 11

US-08-784-582-70
; Sequence 70, Application US/08784582

Patent No. 6110707
GENERAL INFORMATION:

Query Match 26.7%; Score 60.6; DB 3; Length 1884;
Best Local Similarity 84.2%; Pred. No. 5.9e-10;
Matches 80; Conservative 0; Mismatches 14; Indels 1

77 CAGCACAGGCCAATGCCCGTCCTTCCCTGCAGGA 111
||||||| + || | ||||| +
878 CAGCACAGGCCAATGCCCGTCCTTGCCCTGCAGAA 912

RESULT 12

US-08-784-582-72 ; Sequence 72, Application US/08784582
; Patent No. 6110707
; GENERAL INFORMATION:

Query Match	26.7%;	Score 60.6;	DB 3;	Length 2356;
Best Local Similarity	84.2%;	Pred. No. 6.4e-10;		
Matches 80: Conservative		0: Mismatches 14.	Indels 1	

QY 77 CAGCACAGGCCAATGCCGTCCTTCCCCCTGCAGGA 111
||||||| | | | | | | | | |
Db 878 CAGCACAGGCCAATGCCGTCCTTGCCCTGCAGAA 912

RESULTS 13

US-08-068-754-3/c
; Sequence 3, Application US/08068754
; Patent No. 6023011
; GENERAL INFORMATION:
; APPLICANT: Vale, Wyle W.
; APPLICANT: Stenzel-Poore, Mary

APPLICANT: Koob, George F.
APPLICANT: Heinrichs, Stephen C.
TITLE OF INVENTION: CORTICOTROPIN-RELEASING FACTOR
TITLE OF INVENTION: OVERPRODUCING TRANSGENIC MICE
NUMBER OF SEQUENCES: 3
CORRESPONDENCE ADDRESS:
ADDRESSEE: Pretty, Schroeder, Brueggemann & Clark
STREET: 444 South Flower Street, Suite 2000
CITY: Los Angeles
STATE: CA
COUNTRY: USA
ZIP: 90071
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/068,754
FILING DATE: 28-MAY-1993
CLASSIFICATION: 800
ATTORNEY/AGENT INFORMATION:
NAME: Reiter, Stephen E.
REGISTRATION NUMBER: 31,192
REFERENCE/DOCKET NUMBER: P41 9397
TELEPHONE: 619-546-9392
TELEFAX: 619-546-4737
INFORMATION FOR SEQ ID NO: 3:
SEQUENCE CHARACTERISTICS:
LENGTH: 42 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: Other nucleic acid;
DESCRIPTION: Oligonucleotide
US-08-068-754-3

Query Match 18.5%; Score 42; DB 3; Length 42;
Best Local Similarity 100.0%; Pred. No. 0.0001;
Matches 42; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 135 CCCTGGAAGTTGCCACTCCAGTGCACCCAGCCTTGCTCTAA 176
|||||
Db 42 CCCTGGAAGTTGCCACTCCAGTGCACCCAGCCTTGCTCTAA 1

RESULT 14
US-09-325-926-3/c
; Sequence 3, Application US/09325926
; Patent No. 6166287
; GENERAL INFORMATION:
; APPLICANT: Vale, Wyle W.
; Stenzel-Poore, Mary P.
; Koob, George F.
; Heinrichs, Stephen C.
; TITLE OF INVENTION: CORTICOTROPIN-RELEASING FACTOR
; TITLE OF INVENTION: OVERPRODUCING TRANSGENIC MICE
; NUMBER OF SEQUENCES: 3
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Pretty, Schroeder, Brueggemann & Clark
; STREET: 444 South Flower Street, Suite 2000
; CITY: Los Angeles
; STATE: CA
; COUNTRY: USA
; ZIP: 90071
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/09/325,926
FILING DATE: 04-Jun-1999
CLASSIFICATION: <Unknown>
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/068,754
FILING DATE: <Unknown>
ATTORNEY/AGENT INFORMATION:
NAME: Reiter, Stephen E.
REGISTRATION NUMBER: 31,192
REFERENCE/DOCKET NUMBER: P41 9397
TELEPHONE: 619-546-9392
TELEFAX: 619-546-4737
INFORMATION FOR SEQ ID NO: 3:
SEQUENCE CHARACTERISTICS:
LENGTH: 42 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: Other nucleic acid;
DESCRIPTION: Oligonucleotide
SEQUENCE DESCRIPTION: SEQ ID NO: 3:
US-09-325-926-3

Query Match 18.5%; Score 42; DB 4; Length 42;
Best Local Similarity 100.0%; Pred. No. 0.0001;
Matches 42; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 135 CCCTGGAAGTTGCCACTCCAGTGCACCCAGCCTTGCTCTAA 176
|||||
Db 42 CCCTGGAAGTTGCCACTCCAGTGCACCCAGCCTTGCTCTAA 1

RESULT 15
US-08-453-024-10/c
; Sequence 10, Application US/08453024
; Patent No. 5958707
; GENERAL INFORMATION:
; APPLICANT: de Vries, Jan
; APPLICANT: Jenh, Chung-Her
; APPLICANT: Narula, Satwant
; APPLICANT: Zavodny, Paul
; TITLE OF INVENTION: Human Interleukin-4 Antagonist/Agonist Screens
; NUMBER OF SEQUENCES: 10
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Schering-Plough Corporation
; STREET: One Giralda Farms
; CITY: Madison
; STATE: New Jersey
; COUNTRY: USA
; ZIP: 07940
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: Apple Macintosh
; OPERATING SYSTEM: Macintosh 6.0.5
; SOFTWARE: Microsoft Word 5.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/453,024
; FILING DATE: 30-MAY-1995
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/070,162
; FILING DATE: 28-MAY-1993
; APPLICATION NUMBER: US07/869914
; FILING DATE: 16-APR-1992
; APPLICATION NUMBER: US07/770081
; FILING DATE: 03-OCT-1991
; ATTORNEY/AGENT INFORMATION:
; NAME: Dulak, No. 5958707man C.
; REGISTRATION NUMBER: 31,608
; REFERENCE/DOCKET NUMBER: JB0214K
; TELECOMMUNICATION INFORMATION:

```

; TELEPHONE: 201 822 7375
; TELEFAX: 201 822 7039
; TELEX: 219165
; INFORMATION FOR SEQ ID NO: 10:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 51 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
US-08-453-024-10

```

```

Query Match      18.5%  Score 42;  DB 2;  Length 51;
Best Local Similarity 100.0%;  Pred. No. 0.00011;
Matches 42;  Conservative 0;  Mismatches 0;  Indels 0;  Gaps 0;

QY  165 GCCTTGCTCCTAATAAATTAAAGTTGCATCATTTTGTCTGACT 206
      ||||||||||||||||||||||||||||||||||||||||
Db   51  GCCTTGCTCCTAATAAATTAAAGTTGCATCATTTTGTCTGACT 10

```

Search completed: January 17, 2002, 11:52:06
Job time: 17253 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: January 17, 2002, 11:48:13 ; Search time 10436.3 Seconds
(without alignments)
17.388 Million cell updates/sec

Title: US-09-242-202A-10_COPY_1_11
Perfect score: 11
Sequence: 1 GCCTTAAGGC 11

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 1472140 seqs, 8248589755 residues
Total number of hits satisfying chosen parameters: 2944280

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : GenEmbl.*

- 1: gb_ba.*
- 2: gb_htg.*
- 3: gb_in.*
- 4: gb_om.*
- 5: gb_ov.*
- 6: gb_pat.*
- 7: gb_ph.*
- 8: gb_pl.*
- 9: gb_pr.*
- 10: gb_ro.*
- 11: gb_sts.*
- 12: gb_sy.*
- 13: gb_un.*
- 14: gb_vi.*
- 15: em_ba.*
- 16: em_fun.*
- 17: em_hum.*
- 18: em_in.*
- 19: em_om.*
- 20: em_or.*
- 21: em_ov.*
- 22: em_pat.*
- 23: em_ph.*
- 24: em_pl.*
- 25: em_ro.*
- 26: em_sts.*
- 27: em_sy.*
- 28: em_un.*
- 29: em_vi.*
- 30: em_htgo_hum.*
- 31: em_htgo_inv.*
- 32: em_htgo_rod.*
- 33: em_htg_hum.*
- 34: em_htg_inv.*
- 35: em_htg_rod.*
- 36: em_htg_other.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	11	100.0	17	6	AX088916	Sequence
2	11	100.0	33	6	AR069952	Sequence
3	11	100.0	33	6	AX088922	Sequence
4	11	100.0	34	6	AR069934	Sequence
5	11	100.0	34	6	AR076077	Sequence
6	11	100.0	36	6	AR069933	Sequence
7	11	100.0	36	6	AR076076	Sequence
8	11	100.0	37	6	AR075081	Sequence
9	11	100.0	40	6	AR069956	Sequence
10	11	100.0	40	6	AR076090	Sequence
11	11	100.0	42	6	AR069950	Sequence
12	11	100.0	42	6	AR075083	Sequence
13	11	100.0	42	6	AR076083	Sequence
14	11	100.0	48	6	AR069955	Sequence
15	11	100.0	48	6	AR076089	Sequence
16	11	100.0	54	6	AR075075	Sequence
17	11	100.0	54	6	AX088904	Sequence
18	11	100.0	57	6	AR076086	Sequence
19	11	100.0	58	6	AR075077	Sequence
20	11	100.0	58	6	AR076085	Sequence
21	11	100.0	58	6	AX088906	Sequence
22	11	100.0	184	8	MSZ99636	Medicago sa
23	11	100.0	185	6	AX088919	Sequence
24	11	100.0	185	6	AX088920	Sequence
25	11	100.0	185	6	AX088921	Sequence
26	11	100.0	243	11	G05653	human STS W
27	11	100.0	268	11	G04304	human STS W
28	11	100.0	271	11	G64969	FBNI-25 Ran
29	11	100.0	274	1	AB000696	Unidentif
30	11	100.0	275	1	AB000692	Unidentif
31	11	100.0	275	1	AB000693	Unidentif
32	11	100.0	275	1	AB000694	Unidentif
33	11	100.0	276	1	AB000695	Unidentif
34	11	100.0	349	11	G06533	human STS W
35	11	100.0	381	6	AX070615	Sequence
36	11	100.0	382	6	AX026147	Sequence
37	11	100.0	400	11	G14577	human STS S
38	11	100.0	407	11	HS22023S	H.sapiens
39	11	100.0	424	11	HS290153	Homo sapi
40	11	100.0	425	11	HS290145	Homo sapi
41	11	100.0	432	3	DROTNR82	D.melanogas
42	11	100.0	435	14	ESRVLTRL5	Endogenous
43	11	100.0	435	14	ESRVLTRL6	Endogenous
44	11	100.0	436	3	DROTNR81	D. melanoga
45	11	100.0	442	14	ENJRVLTR3	Endogenous

ALIGNMENTS

RESULT 1	AX088916	Sequence	17 bp	DNA	PAT	17-MAR-2001
LOCUS	AX088916	Sequence	17 from Patent	WO0114534.		
DEFINITION	AX088916	Sequence	17 from Patent	WO0114534.		
ACCESSION	AX088916	Sequence	17 from Patent	WO0114534.		
VERSION	AX088916.1	GI:13397675				
KEYWORDS						
SOURCE		Bacillus subtilis.				
ORGANISM		Bacillus subtilis				

REFERENCE
AUTHORS Thomas,M.D. and Brown,K.M.
TITLE Polypeptides having pectin acetyltransferase activity and nucleic acids encoding same
JOURNAL Patent: WO 0114534-A 17 01-MAR-2001;
FEATURES Novozymes Biotech, Inc. (US)
source Location/Qualifiers
1. .17
/organism="Bacillus subtilis"

```
BASE COUNT      3 a      5 c      6 g      3 t
ORIGIN

Query Match      100.0%; Score 11; DB 6; Length 17;
Best Local Similarity 100.0%; Pred. No. 1.2e+04;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY  1 GCCTTAAGGCG 11
    |||||
Db  2 GCCTTAAGGCG 12

RESULT 4
AR069934
LOCUS      AR069934      34 bp      DNA
DEFINITION Sequence 19 from patent US 5891701.
ACCESSION  AR069934
VERSION     AR069934.1 GI:7220822
KEYWORDS
SOURCE      Unknown.
ORGANISM     Unknown.
REFERENCE    1 (bases 1 to 34)
AUTHORS      Sloma,A. and Christianson,L.
TITLE        Nucleic acids encoding a polypeptide having protease activity.
JOURNAL      Patent: US 5891701-A 19 06-APR-1999;
FEATURES     Location/Qualifiers
             source
             1..34
             /organism="unknown"
BASE COUNT   5 a      10 c      11 g      8 t
ORIGIN

Query Match      100.0%; Score 11; DB 6; Length 34;
Best Local Similarity 100.0%; Pred. No. 1.1e+04;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY  1 GCCTTAAGGCG 11
    |||||
Db  3 GCCTTAAGGCG 13

RESULT 5
AR076077
LOCUS      AR076077      34 bp      DNA
DEFINITION Sequence 19 from patent US 5958728.
ACCESSION  AR076077
VERSION     AR076077.1 GI:10002823
KEYWORDS
SOURCE      Unknown.
ORGANISM     Unknown.
REFERENCE    1 (bases 1 to 34)
AUTHORS      Sloma,A., Sternberg,D., Adams,L.F. and Brown,S.
TITLE        Methods for producing polypeptides in mutants of bacillus cells
JOURNAL      Patent: US 5958728-A 19 28-SEP-1999;
FEATURES     Location/Qualifiers
             source
             1..34
             /organism="unknown"
BASE COUNT   5 a      10 c      11 g      8 t
ORIGIN

Query Match      100.0%; Score 11; DB 6; Length 34;
Best Local Similarity 100.0%; Pred. No. 1.1e+04;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY  1 GCCTTAAGGCG 11
    |||||
Db  3 GCCTTAAGGCG 13

RESULT 6
AR069933/c
LOCUS      AR069933      36 bp      DNA
DEFINITION Sequence 18 from patent US 5891701.
ACCESSION  AR069933
VERSION     AR069933.1 GI:7220821
KEYWORDS
SOURCE      Unknown.
```

```
BASE COUNT      3 a      5 c      6 g      3 t
ORIGIN

Query Match      100.0%; Score 11; DB 6; Length 17;
Best Local Similarity 100.0%; Pred. No. 1.2e+04;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY  1 GCCTTAAGGCG 11
    |||||
Db  2 GCCTTAAGGCG 12

RESULT 2
AR069952
LOCUS      AR069952      33 bp      DNA
DEFINITION Sequence 39 from patent US 5891701.
ACCESSION  AR069952
VERSION     AR069952.1 GI:7220840
KEYWORDS
SOURCE      Unknown.
ORGANISM     Unknown.
REFERENCE    1 (bases 1 to 33)
AUTHORS      Sloma,A. and Christianson,L.
TITLE        Nucleic acids encoding a polypeptide having protease activity
JOURNAL      Patent: US 5891701-A 39 06-APR-1999;
FEATURES     Location/Qualifiers
             source
             1..33
             /organism="unknown"
BASE COUNT   4 a      11 c      9 g      9 t
ORIGIN

Query Match      100.0%; Score 11; DB 6; Length 33;
Best Local Similarity 100.0%; Pred. No. 1.1e+04;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY  1 GCCTTAAGGCG 11
    |||||
Db  5 GCCTTAAGGCG 15

RESULT 3
AX088922
LOCUS      AX088922      33 bp      DNA
DEFINITION Sequence 23 from Patent WO0114534.
ACCESSION  AX088922
VERSION     AX088922.1 GI:13397681
KEYWORDS
SOURCE      Bacillus subtilis.
ORGANISM     Bacillus subtilis
REFERENCE    1 (bases 1 to 33)
AUTHORS      Thomas,M.D. and Brown,K.M.
TITLE        Polypeptides having pectin acetylesterase activity and nucleic
JOURNAL      Patent: WO 0114534-A 23 01-MAR-2001;
FEATURES     Location/Qualifiers
             source
             1..33
             /organism="Bacillus subtilis"
BASE COUNT   4 a      11 c      10 g      8 t
ORIGIN

Query Match      100.0%; Score 11; DB 6; Length 33;
Best Local Similarity 100.0%; Pred. No. 1.1e+04;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```


ORGANISM Unknown.
Unclassified.
REFERENCE 1 (bases 1 to 36)
AUTHORS Sloma,A. and Christianson,L.
TITLE Nucleic acids encoding a polypeptide having protease activity
JOURNAL Patent: US 5891701-A 18 06-APR-1999;
FEATURES Location/Qualifiers
source 1..36
BASE COUNT 11 a 12 c 10 g 3 t
ORIGIN

Query Match
Best Local Similarity 100.0%; Score 11; DB 6; Length 36;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GCCTTAAGGC 11
Db 22 GCCTTAAGGC 12

RESULT 7
AR076076/c
LOCUS AR076076 36 bp DNA PAT 30-AUG-2000
DEFINITION Sequence 18 from patent US 5958728.
ACCESSION AR076076
VERSION AR076076.1 GI:10002822
KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.
Unclassified.
REFERENCE 1 (bases 1 to 36)
AUTHORS Sloma,A., Sternberg,D., Adams,L.F. and Brown,S.
TITLE Methods for producing polypeptides in mutants of bacillus cells
JOURNAL Patent: US 5958728-A 18 28-SEP-1999;
FEATURES Location/Qualifiers
source 1..36
BASE COUNT 11 a 12 c 10 g 3 t
ORIGIN

Query Match
Best Local Similarity 100.0%; Score 11; DB 6; Length 36;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GCCTTAAGGC 11
Db 22 GCCTTAAGGC 12

RESULT 8
AR075081
LOCUS AR075081 37 bp DNA PAT 28-AUG-2000
DEFINITION Sequence 7 from patent US 5955310.
ACCESSION AR075081
VERSION AR075081.1 GI:10001833
KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.
Unclassified.
REFERENCE 1 (bases 1 to 37)
AUTHORS Widner,W., Sloma,A. and Thomas,M.D.
TITLE Methods for producing a polypeptide in a bacillus cell
JOURNAL Patent: US 5955310-A 7 21-SEP-1999;
FEATURES Location/Qualifiers
source 1..37
BASE COUNT 4 a 12 c 10 g 11 t
ORIGIN

Query Match
Best Local Similarity 100.0%; Score 11; DB 6; Length 37;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GCCTTAAGGC 11
Db 5 GCCTTAAGGC 15

RESULT 9
AR069956/c
LOCUS AR069956 40 bp DNA PAT 18-FEB-2000
DEFINITION Sequence 46 from patent US 5891701.
ACCESSION AR069956
VERSION AR069956.1 GI:7220844
KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.
Unclassified.
REFERENCE 1 (bases 1 to 40)
AUTHORS Sloma,A. and Christianson,L.
TITLE Nucleic acids encoding a polypeptide having protease activity
JOURNAL Patent: US 5891701-A 46 06-APR-1999;
FEATURES Location/Qualifiers
source 1..40
BASE COUNT 7 a 14 c 13 g 6 t
ORIGIN

Query Match
Best Local Similarity 100.0%; Score 11; DB 6; Length 40;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GCCTTAAGGC 11
Db 38 GCCTTAAGGC 28

RESULT 10
AR076090/c
LOCUS AR076090 40 bp DNA PAT 30-AUG-2000
DEFINITION Sequence 32 from patent US 5958728.
ACCESSION AR076090
VERSION AR076090.1 GI:10002836
KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.
Unclassified.
REFERENCE 1 (bases 1 to 40)
AUTHORS Sloma,A., Sternberg,D., Adams,L.F. and Brown,S.
TITLE Methods for producing polypeptides in mutants of bacillus cells
JOURNAL Patent: US 5958728-A 32 28-SEP-1999;
FEATURES Location/Qualifiers
source 1..40
BASE COUNT 7 a 14 c 13 g 6 t
ORIGIN

Query Match
Best Local Similarity 100.0%; Score 11; DB 6; Length 40;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GCCTTAAGGC 11
Db 38 GCCTTAAGGC 28

RESULT 11
AR069950
LOCUS AR069950 42 bp DNA PAT 18-FEB-2000
DEFINITION Sequence 37 from patent US 5891701.

[illegible]

Fri Jan 18 08:27:34 2002

us-09-242-202a-10_copy_1_11.rge

Page 5

Search completed: January 17, 2002, 11:48:13
Job time: 17105 sec

THIS PAGE BLANK (USPTO)

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: January 17, 2002, 12:02:18 ; Search time 599.86 Seconds
(without alignments)
15.721 Million cell updates/sec

Title: US-09-242-202a-10_COPY_1_11

Perfect score: 11

Sequence: 1 GCCTTAGGGC 11

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 930621 seqs, 428662619 residues

Total number of hits satisfying chosen parameters: 1861242

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

N_Geneseq_1101:*

- 1: /SIDS2/gcgdata/geneseq/geneseq/NA1980.DAT:*
- 2: /SIDS2/gcgdata/geneseq/geneseq/NA1981.DAT:*
- 3: /SIDS2/gcgdata/geneseq/geneseq/NA1982.DAT:*
- 4: /SIDS2/gcgdata/geneseq/geneseq/NA1983.DAT:*
- 5: /SIDS2/gcgdata/geneseq/geneseq/NA1984.DAT:*
- 6: /SIDS2/gcgdata/geneseq/geneseq/NA1985.DAT:*
- 7: /SIDS2/gcgdata/geneseq/geneseq/NA1986.DAT:*
- 8: /SIDS2/gcgdata/geneseq/geneseq/NA1987.DAT:*
- 9: /SIDS2/gcgdata/geneseq/geneseq/NA1988.DAT:*
- 10: /SIDS2/gcgdata/geneseq/geneseq/NA1989.DAT:*
- 11: /SIDS2/gcgdata/geneseq/geneseq/NA1990.DAT:*
- 12: /SIDS2/gcgdata/geneseq/geneseq/NA1991.DAT:*
- 13: /SIDS2/gcgdata/geneseq/geneseq/NA1992.DAT:*
- 14: /SIDS2/gcgdata/geneseq/geneseq/NA1993.DAT:*
- 15: /SIDS2/gcgdata/geneseq/geneseq/NA1994.DAT:*
- 16: /SIDS2/gcgdata/geneseq/geneseq/NA1995.DAT:*
- 17: /SIDS2/gcgdata/geneseq/geneseq/NA1996.DAT:*
- 18: /SIDS2/gcgdata/geneseq/geneseq/NA1997.DAT:*
- 19: /SIDS2/gcgdata/geneseq/geneseq/NA1998.DAT:*
- 20: /SIDS2/gcgdata/geneseq/geneseq/NA1999.DAT:*
- 21: /SIDS2/gcgdata/geneseq/geneseq/NA2000.DAT:*
- 22: /SIDS2/gcgdata/geneseq/geneseq/NA2001.DAT:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	11	100.0	11	19	AAV21736 Humanised vector 3
2	11	100.0	17	20	AAZ23342 Bacillus amyloliqu
3	11	100.0	17	22	AAD09919 PCR primer #1 for
4	11	100.0	17	22	AAF62642 amyQ promoter olig
5	11	100.0	33	20	AAZ23344 Bacillus amyloliqu
6	11	100.0	33	20	AAZ23348 Bacillus amyloliqu
7	11	100.0	33	20	AAV82419 Bacillus amyL prom
8	11	100.0	33	22	AAD09925 Short consensus am
9	11	100.0	33	22	AAD09928 amyQ promoter olig
10	11	100.0	33	22	AAF62648 PCR primer 19 used
11	11	100.0	34	19	AAV38595

12	11	100.0	34	20	AAV82401	Bacillus subtilis
13	11	100.0	36	19	AAV38594	PCR primer 18 used
14	11	100.0	36	20	AAV82400	Bacillus subtilis
15	11	100.0	37	20	AAZ23328	Bacillus lichenifo
16	11	100.0	37	22	AAD09904	Bacillus lichenifo
17	11	100.0	42	19	AAV38601	PCR primer 25 used
18	11	100.0	42	20	AAZ23330	Bacillus amyloliqu
19	11	100.0	42	20	AAV82417	Bacillus amyQ prom
20	11	100.0	42	22	AAD09906	Bacillus amyloliqu
21	11	100.0	51	22	AAD09899	Synthetic polylink
22	11	100.0	54	22	AAD09898	Synthetic polylink
23	11	100.0	54	22	AAF62630	Polylinker #1. Sy
24	11	100.0	58	22	AAD09900	Synthetic polylink
25	11	100.0	58	22	AAD09901	Synthetic polylink
26	11	100.0	58	22	AAF62632	Polylinker #3. Sy
27	11	100.0	63	20	AAZ34335	HIV polyA(2) compl
28	11	100.0	65	20	AAZ34334	HIV polyA(1) prime
29	11	100.0	185	20	AAZ23322	B. thuringiensis w
30	11	100.0	185	20	AAZ23323	B. amyloliquefacie
31	11	100.0	185	20	AAZ23324	B. amyloliquefacie
32	11	100.0	185	20	AAZ23325	Bacillus sp. amyQ
33	11	100.0	185	22	AAD09922	Bacillus amyloliqu
34	11	100.0	185	22	AAD09923	Bacillus amyloliqu
35	11	100.0	185	22	AAD09924	Bacillus amyloliqu
36	11	100.0	185	22	AAD09931	Bacillus amyloliqu
37	11	100.0	185	22	AAF62645	Wild type amyQ pro
38	11	100.0	185	22	AAF62646	Consensus amyQ pro
39	11	100.0	185	22	AAF62647	Consensus amyQ pro
40	11	100.0	227	19	AAV21719	Combined 3' splice
41	11	100.0	228	19	AAV21723	Humanised vector s
42	11	100.0	252	19	AAV21720	Humanised polynuc
43	11	100.0	300	20	AAZ13626	Human gene express
44	11	100.0	308	21	AAZ32016	Plant microsatelli
45	11	100.0	338	21	AAZ31512	Plant microsatelli

ALIGNMENTS

RESULT 1

AAV21736

ID AAV21736 standard; cDNA; 11 BP.

XX

AC AAV21736;

XX

DT 17-AUG-1998 (first entry)

XX

DE Humanised vector 3' sequence acceptance site.

XX

KW Vector; vaccine; tumour; antigen; plasmid pITL; ss.

XX

OS Synthetic.

XX

PN WO9806863-A1.

XX

PD 19-FEB-1998.

XX

PF 14-AUG-1997; 97WO-US14306.

XX

PR 14-AUG-1996; 96US-0023931.

XX

PA (USSH) US DEPT HEALTH & HUMAN SERVICES.

XX

PI Nelson EL, Nelson PJ;

XX

DR WPI; 1998-159552/14.

XX

PT Humanised polynucleotide vectors - comprising human derived promoter and sequence acceptance site, used for the production of vaccines

XX

PS Claim 13; Page 104; 125pp; English.

XX

CC This DNA sequence comprises a 5' acceptance site of novel humanised

CC polynucleotide vectors such as pITL (see AAV21724). Such vectors
 CC comprise a human-derived promoter or mammalian homologue and a
 CC sequence acceptance site that is designed to directionally accept
 CC sequence specific products from RT-PCR based cloning strategies
 CC via a unique site within an interrupted palindrome recognition
 CC sequence for a restriction endonuclease which is incorporated into
 CC the PCR primer. In this embodiment, the palindrome recognition
 CC sequence is for BglI. The 5' acceptance site is given in AAV21735.
 CC The novel vectors are used to express target antigens, especially
 CC tumour antigens. They are non-replicating in mammalian cells but
 CC are capable of extended stable expression of target sequences, and
 CC generate immune responses to the target sequences with little or no
 CC immune response to the other vector components.

XX SQ Sequence 11 BP; 2 A; 3 C; 4 G; 2 T; 0 other;

Query Match 100.0%; Score 11; DB 19; Length 11;
 Best Local Similarity 100.0%; Pred. No. 3e+02; Indels 0; Gaps 0;
 Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GCCTTAAGGCG 11
 Db 1 gccttaaggcg 11
 |||||

RESULT 2
 AAZ23342
 ID AAZ23342 standard; DNA; 17 BP.
 XX
 AC AAZ23342;
 XX
 DT 06-DEC-1999 (first entry)
 XX
 DE Bacillus amyloliquefaciens consensus amyQ gene promoter PCR primer 2.
 XX
 KW Tandem promoter; consensus promoter; enzyme production; hormone;
 KW antibody; reporter; marker gene; cell regulation; PCR primer; amyQ;
 KW alpha-amylase; ss.
 OS Synthetic.
 OS Bacillus amyloliquefaciens.
 XX
 PN WO9943835-A2.
 XX
 XX 02-SEP-1999.
 XX
 PF 26-FEB-1999; 99WO-US04360.
 XX
 PR 26-FEB-1999; 98US-0031442.
 XX
 PA (NOVO) NOVO NORDISK BIOTECH INC.
 XX
 PI Widner W, Sloma A, Thomas MD;
 XX
 DR WPI; 1999-561370/47.
 XX
 PT Production of polypeptide in Bacillus using specific promoters,
 PT particularly for producing enzymes -
 XX
 PS Example 18; Page 35; 90pp; English.

CC This invention describes a novel method for the production of a
 CC polypeptide in Bacillus using specific tandem or consensus promoters.
 CC The method is used to produce homologous or particularly heterologous
 CC proteins, particularly enzymes (specifically serine protease, maltogenic
 CC alpha-amylase and pullulanase), but also hormones, antibodies, reporters
 CC etc. The specified promoters provide increased expression of the sequence
 CC which encodes the polypeptide of the invention. After incorporation of
 CC the nucleic acid construct of the invention, any marker gene may be
 CC deleted, resulting in a cell that is preferred for environmental and
 CC regulatory regions. This sequence represents a PCR primer used to amplify
 CC the Bacillus amyloliquefaciens consensus amyQ promoter.

XX SQ Sequence 17 BP; 3 A; 5 C; 6 G; 3 T; 0 other;

Query Match 100.0%; Score 11; DB 20; Length 17;
 Best Local Similarity 100.0%; Pred. No. 3.2e+02; Indels 0; Gaps 0;
 Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GCCTTAAGGCG 11
 Db 2 gccttaaggcg 12
 |||||

RESULT 3
 AAD09919
 ID AAD09919 standard; DNA; 17 BP.
 XX
 AC AAD09919;
 XX
 DT 12-SEP-2001 (first entry)
 XX
 DE PCR primer #1 for constructing mutated consensus amyQ promoter.
 XX
 KW Bacillus cell; tandem promoter; consensus promoter;
 KW polypeptide production; amyL promoter; amyQ promoter; aprH promoter;
 KW cryIIIA promoter; subtilisin Carlsberg gene promoter; PCR primer; ss.
 XX
 OS Bacillus amyloliquefaciens.
 XX
 PN US6255076-B1.
 XX
 PD 03-JUL-2001.
 XX
 PF 26-FEB-1999; 99US-0258377.
 XX
 PR 26-FEB-1999; 98US-0031442.
 XX
 PA (NOVO-) NOVOZYMES BIOTECH INC.
 XX
 PI Widner W, Sloma A, Thomas MD;
 XX
 DR WPI; 2001-440518/47.
 XX
 PT Producing a polypeptide in a Bacillus strain comprises cultivating a
 PT Bacillus cell, which contains a nucleic acid construct comprising a
 PT tandem promoter and consensus promoters
 XX
 PS Example 18; Column 25; 54pp; English.

CC The present invention relates to a method for producing a polypeptide,
 CC comprises cultivating a Bacillus cell, which contains a nucleic acid
 CC construct comprising a tandem promoter and consensus promoters. The
 CC Bacillus is cultivated in a medium conducive for the production of the
 CC polypeptide. Each promoter sequence of the tandem promoter is operably
 CC linked to a nucleic acid sequence encoding the polypeptide. The tandem
 CC promoter comprises amyL promoter, amyQ promoter, aprH promoter, cryIIIA
 CC promoter or subtilisin Carlsberg gene promoter. The consensus promoters
 CC of the amyL promoter, amyQ promoter, aprH promoter, cryIIIA promoter
 CC or subtilisin Carlsberg gene promoter have the sequence TTGACA for
 CC the -35 region and TATAAT for the -10 region. The method further
 CC comprises isolating the polypeptide from the cultivation medium. The
 CC method is useful for producing a polypeptide in a Bacillus strain.
 CC The present sequence is a PCR primer used in the construction of a
 CC consensus mutated Bacillus amyloliquefaciens alpha-amylase (amyQ)
 CC promoter. The primer generates a double-stranded fragment comprising
 CC 137 bp of the upstream region of the amyQ promoter, used in the
 CC exemplification of the invention.

XX SQ Sequence 17 BP; 3 A; 5 C; 6 G; 3 T; 0 other;

Query Match 100.0%; Score 11; DB 22; Length 17;
 Best Local Similarity 100.0%; Pred. No. 3.2e+02;

```
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GCCTTAAGGCG 11
Db 2 gccttaaggcg 12

RESULT 4
AAF62642
ID AAF62642 standard; DNA; 17 BP.
XX
AC AAF62642;
XX
DT 03-MAY-2001 (first entry)
XX
DE amyQ promoter oligonucleotide #3.
XX
KW Pectin acylesterase; degrade; plant cell wall; ss.
XX
OS Synthetic.
XX
PN US6184028-B1.
XX
PD 06-FEB-2001.
XX
PF 26-AUG-1999; 99US-0384305.
XX
PR 26-AUG-1999; 99US-0384305.
XX
PA (NOVO ) NOVO NORDISK BIOTECH INC.
XX
PI Thomas MD, Brown KM;
XX
XX WPI; 2001-190946/19.
XX
PT Novel isolated polypeptide having pectin acylesterase activity useful
PT for degrading pectic substances and in degradation or modification of
PT acetylated pectins and plant cell walls
XX
XX Example 9; Column 28; 35pp; English.
XX
CC The present invention relates to Bacillus subtilis pectin
CC acylesterase protein. The invention is useful for degrading
CC a pectic substance. It is also useful for degrading soluble and
CC insoluble pectins with varying degrees of esterification,
CC clarification etc. The protein may be used alone or in
CC combination with other enzymes for the degradation or modification
CC of acetylated pectins, degradation or modification of plant cell walls.
XX
SQ Sequence 17 BP; 3 A; 5 C; 6 G; 3 T; 0 Other;

Query Match 100.0%; Score 11; DB 22; Length 17;
Best Local Similarity 100.0%; Pred. NO. 3.2e+02;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GCCTTAAGGCG 11
Db 2 gccttaaggcg 12

RESULT 5
AA223344
ID AA223344 standard; DNA; 33 BP.
XX
AC AA223344;
XX
DT 06-DEC-1999 (first entry)
XX
DE Bacillus amyloliquefaciens short consensus amyQ promoter PCR primer 1.
XX
KW Tandem promoter; consensus promoter; enzyme production; hormone;
KW antibody; reporter; marker gene; cell regulation; PCR primer; amyQ;

Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GCCTTAAGGCG 11
Db 2 gccttaaggcg 12

RESULT 6
AA223348
ID AA223348 standard; DNA; 33 BP.
XX
AC AA223348;
XX
DT 06-DEC-1999 (first entry)
XX
DE Bacillus amyloliquefaciens short consensus amyQ promoter PCR primer 5.
XX
KW Tandem promoter; consensus promoter; enzyme production; hormone;
KW antibody; reporter; marker gene; cell regulation; PCR primer; amyQ;
KW alpha-amylase; SAVINASE; ss.
XX
OS Synthetic.
OS Bacillus amyloliquefaciens.
XX
PN WO9943835-A2.
XX
PD 02-SEP-1999.
XX
PF 26-FEB-1999; 99WO-US04360.
XX
```

```
KW alpha-amylase; SAVINASE; ss.
XX
OS Synthetic.
OS Bacillus amyloliquefaciens.
XX
PN WO9943835-A2.
XX
PD 02-SEP-1999.
XX
PF 26-FEB-1999; 99WO-US04360.
XX
PR 26-FEB-1998; 98US-0031442.
XX
PA (NOVO ) NOVO NORDISK BIOTECH INC.
XX
PI Widner W, Sloma A, Thomas MD;
XX
XX WPI; 1999-561370/47.
DR Production of polypeptide in Bacillus using specific promoters,
XX particularly for producing enzymes
XX
PS Example 19; Page 36; 90pp; English.
XX
CC This invention describes a novel method for the production of a
CC polypeptide in Bacillus using specific tandem or consensus promoters.
CC The method is used to produce homologous or particularly heterologous
CC proteins, particularly enzymes (specifically serine protease, maltogenic
CC alpha-amylase and pullulanase), but also hormones, antibodies, reporters
CC etc. The specified promoters provide increased expression of the sequence
CC which encodes the polypeptide of the invention. After incorporation of
CC the nucleic acid construct of the invention, any marker gene may be
CC deleted, resulting in a cell that is preferred for environmental and
CC regulatory regions. This sequence represents a PCR primer used to amplify
CC the Bacillus amyloliquefaciens short consensus amyQ promoter which is
CC used in the construction of a short consensus amyQ promoter-SAVINASE(RTM)
CC gene expression cassette.
XX
SQ Sequence 33 BP; 4 A; 11 C; 10 G; 8 T; 0 Other;
```

```
Query Match 100.0%; Score 11; DB 20; Length 33;
Best Local Similarity 100.0%; Pred. NO. 3.4e+02;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GCCTTAAGGCG 11
Db 2 gccttaaggcg 12

RESULT 6
AA223348
ID AA223348 standard; DNA; 33 BP.
XX
AC AA223348;
XX
DT 06-DEC-1999 (first entry)
XX
DE Bacillus amyloliquefaciens short consensus amyQ promoter PCR primer 5.
XX
KW Tandem promoter; consensus promoter; enzyme production; hormone;
KW antibody; reporter; marker gene; cell regulation; PCR primer; amyQ;
KW alpha-amylase; SAVINASE; ss.
XX
OS Synthetic.
OS Bacillus amyloliquefaciens.
XX
PN WO9943835-A2.
XX
PD 02-SEP-1999.
XX
PF 26-FEB-1999; 99WO-US04360.
XX
```

PR 26-FEB-1998; 98US-0031442.
 XX (NOVO) NOVO NORDISK BIOTECH INC.
 XX Widner W, Sloma A, Thomas MD;
 XX WPI; 1999-561370/47.
 XX Production of polypeptide in *Bacillus* using specific promoters,
 PT particularly for producing enzymes -
 XX Example 25; Page 41; 90pp; English.
 XX This invention describes a novel method for the production of a
 CC polypeptide in *Bacillus* using specific tandem or consensus promoters.
 CC The method is used to produce homologous or particularly heterologous
 CC proteins, particularly enzymes (specifically serine protease, maltogenic
 CC alpha-amylase and pullulanase), but also hormones, antibodies, reporters
 CC etc. The specified promoters provide increased expression of the sequence
 CC which encodes the polypeptide of the invention. After incorporation of
 CC the nucleic acid construct of the invention, any marker gene may be
 CC deleted, resulting in a cell that is preferred for environmental and
 CC regulatory regions. This sequence represents a PCR primer used to amplify
 CC the *Bacillus amyloliquefaciens* short consensus amyQ trimer promoter, which
 CC is used in the construction of a short consensus amyQ dimer promoter.
 CC SAVINASE(RTM) gene expression cassette.
 XX Sequence 33 BP; 4 A; 11 C; 10 G; 8 T; 0 other;
 SQ

Query Match 100.0%; Score 11; DB 20; Length 33;
 Best Local Similarity 100.0%; Pred. No. 3.4e+02;
 Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GCCTTAAGGCG 11
 DB 2 gccttaaggcg 12
 |||||

RESULT 7
 ID AAV82419 standard; DNA; 33 BP.
 XX AAV82419;
 XX 12-APR-1999 (first entry)
 XX *Bacillus amyl* promoter PCR primer term1sfi.
 DE Protease; detergent; surfactant; leather processing; debittering;
 KW flavour; PCR; primer; amyL; amylase; promoter; ss.
 OS Synthetic.
 XX *Bacillus amyloliquefaciens*.
 PN W09856927-A2.
 XX 17-DEC-1998.
 XX 09-JUN-1998; 98WO-US12005.
 XX 12-JUN-1997; 97US-0873479.
 XX (NOVO) NOVO NORDISK BIOTECH INC.
 XX Christianson L, Sloma A;
 XX WPI; 1999-080908/07.
 XX Novel protease from *Bacillus subtilis* LC20 - useful in laundry and
 PT dishwashing detergents and for leather processing
 XX Example 12; Page 36; 77pp; English.

XX Primers term1sfi and 2isfi (see AAV82419-20) were used for the PCR
 CC amplification of the promoter region of the *TERMAMYL* amylase
 CC amyL gene of *Bacillus amyloliquefaciens*. Primer term1sfi was
 CC used to add an SfiI site to the 5' end of the promoter. The
 CC amplified promoter was inserted upstream of a reconstructed
 CC *Bacillus* JP170 protease gene (see AAV82382) in vector pS2882-MCS,
 CC giving plasmid p170TERM. The gene encodes a protease (see AAW89547)
 CC useful in detergent compositions and in leather processing. Host
 CC cells, vectors and methods for recombinant production of the
 CC protease are provided.
 XX Sequence 33 BP; 4 A; 11 C; 9 G; 9 T; 0 other;
 SQ

Query Match 100.0%; Score 11; DB 20; Length 33;
 Best Local Similarity 100.0%; Pred. No. 3.4e+02;
 Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GCCTTAAGGCG 11
 DB 5 gccttaaggcg 15
 |||||

RESULT 8
 ID AAD09925 standard; DNA; 33 BP.
 XX AAD09925;
 XX 12-SEP-2001 (first entry)
 XX Short consensus amyQ promoter amplifying PCR primer.
 DE *Bacillus* cell; tandem promoter; consensus promoter;
 KW polypeptide production; amyL promoter; amyQ promoter; aprH promoter;
 KW cryIIIA promoter; subtilisin Carlsberg gene promoter; PCR primer; ss.
 OS *Bacillus amyloliquefaciens*.
 XX US6255076-B1.
 XX 03-JUL-2001.
 XX 26-FEB-1999; 99US-0258377.
 XX 26-FEB-1998; 98US-0031442.
 XX (NOVO-) NOVOZYMES BIOFTECH INC.
 XX Widner W, Sloma A, Thomas MD;
 XX WPI; 2001-440518/47.
 XX Producing a polypeptide in a *Bacillus* strain comprises cultivating a
 PT *Bacillus* cell, which contains a nucleic acid construct comprising a
 PT tandem promoter and consensus promoters -
 XX Example 19; Column 26; 54pp; English.
 XX The present invention relates to a method for producing a polypeptide,
 CC comprises cultivating a *Bacillus* cell, which contains a nucleic acid
 CC construct comprising a tandem promoter and consensus promoters. The
 CC *Bacillus* is cultivated in a medium conducive for the production of the
 CC polypeptide. Each promoter sequence of the tandem promoter is operably
 CC linked to a nucleic acid sequence encoding the polypeptide. The tandem
 CC promoter comprises amyL promoter, amyQ promoter, aprH promoter, cryIIIA
 CC promoter or subtilisin Carlsberg gene promoter. The consensus promoters
 CC of the amyL promoter, amyQ promoter, aprH promoter, cryIIIA promoter
 CC or subtilisin Carlsberg gene promoter have the sequence TTGACA for
 CC the -35 region and TATAAT for the -10 region. The method further
 CC comprises isolating the polypeptide from the cultivation medium. The
 CC method is useful for producing a polypeptide in a *Bacillus* strain.

CC The present sequence is a PCR primer used in the construction of a
CC short consensus mutated Bacillus amyloliquefaciens alpha-amylase (amyQ)
CC promoter gene expression cassette, used in the exemplification of
CC the invention.
XX
SQ Sequence 33 BP; 4 A; 11 C; 10 G; 8 T; 0 other;

Query Match 100.0%; Score 11; DB 22; Length 33;
Best Local Similarity 100.0%; Pred. No. 3.4e+02;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GCCTTAAGGCG 11
Db 2 gccttaagggc 12
|||||

RESULT 9
AAD09928
ID AAD09928 standard; DNA; 33 BP.
XX AC AAD09928;
XX DT 12-SEP-2001 (first entry)
XX Short consensus amyQ dimer promoter amplifying PCR primer.
XX
XX Bacillus cell; tandem promoter; consensus promoter;
KW polypeptide production; amyL promoter; amyQ promoter; aprH promoter;
KW cryIIIA promoter; subtilisin Carlsberg gene promoter; PCR primer; ss.
XX
XX Bacillus amyloliquefaciens.
XX
XX US6255076-B1.
XX
XX 03-JUL-2001.
XX
XX 26-FEB-1999; 99US-0258377.
XX
XX 26-FEB-1998; 98US-0031442.
XX
XX (NOVO-) NOVOZYMES BIOTECH INC.
XX
XX Widner W, Sloma A, Thomas MD;
XX
XX WPI; 2001-440518/47.
XX

PT Producing a polypeptide in a Bacillus strain comprises cultivating a
PT Bacillus cell, which contains a nucleic acid construct comprising a
PT tandem promoter and consensus promoters -
XX
XX Example 25; Column 30; 54pp; English.

CC The present invention relates to a method for producing a polypeptide,
CC comprises cultivating a Bacillus cell, which contains a nucleic acid
CC construct comprising a tandem promoter and consensus promoters. The
CC Bacillus is cultivated in a medium conducive for the production of the
CC polypeptide. Each promoter sequence of the tandem promoter is operably
CC linked to a nucleic acid sequence encoding the polypeptide. The tandem
CC promoter comprises amyL promoter, amyQ promoter, aprH promoter, cryIIIA
CC promoter or subtilisin Carlsberg gene promoter. The consensus promoters
CC of the amyL promoter, amyQ promoter, aprH promoter, cryIIIA promoter
CC or subtilisin Carlsberg gene promoter have the sequence TTCACA for
CC the -35 region and TATAAT for the -10 region. The method further
CC comprises isolating the polypeptide from the cultivation medium. The
CC method is useful for producing a polypeptide in a Bacillus strain.
CC The present sequence is a PCR primer used in the construction of a
CC short consensus mutated Bacillus amyloliquefaciens alpha-amylase (amyQ)
CC dimer promoter gene expression cassette, used in the exemplification of
CC the invention.

XX Sequence 33 BP; 4 A; 11 C; 10 G; 8 T; 0 other;

Query Match 100.0%; Score 11; DB 22; Length 33;
Best Local Similarity 100.0%; Pred. No. 3.4e+02;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GCCTTAAGGCG 11
Db 2 gccttaagggc 12
|||||

RESULT 10
AAF62648
ID AAF62648 standard; DNA; 33 BP.
XX AC AAF62648;
XX DT 03-MAY-2001 (first entry)
XX amyQ promoter oligonucleotide #6.
DE Pectin acetylsterase; degrade; plant cell wall; ss.
XX
XX Synthetic.
XX
XX US6184028-B1.
XX
XX 06-FEB-2001.
XX
XX 26-AUG-1999; 99US-0384305.
XX
XX 26-AUG-1999; 99US-0384305.
XX
XX (NOVO) NOVO NORDISK BIOTECH INC.
XX
XX Thomas MD, Brown KM;
XX
XX WPI; 2001-190946/19.
XX

PT Novel isolated polypeptide having pectin acetylsterase activity useful
PT for degrading pectic substances and in degradation or modification of
PT acetylated pectins and plant cell walls -
XX
XX Example 10; Column 29; 35pp; English.

CC The present invention relates to Bacillus subtilis pectin
CC acetylsterase protein. The invention is useful for degrading
CC a pectic substance. It is also useful for degrading soluble and
CC insoluble pectins with varying degrees of esterification,
CC clarification etc. The protein may be used alone or in
CC combination with other enzymes for the degradation or modification
CC of acetylated pectins, degradation or modification of plant cell walls.
XX
SQ Sequence 33 BP; 4 A; 11 C; 10 G; 8 T; 0 other;

Query Match 100.0%; Score 11; DB 22; Length 33;
Best Local Similarity 100.0%; Pred. No. 3.4e+02;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GCCTTAAGGCG 11
Db 2 gccttaagggc 12
|||||

RESULT 11
AAV38595
ID AAV38595 standard; DNA; 34 BP.
XX AC AAV38595;
XX DT 17-SEP-1998 (first entry)
XX PCR primer 19 used to amplify a portion of the amyE gene.
DE

XX amyE gene; strain Al64; mutant Bacillus; biosynthesis; surfactin;
 KW production; hormone; enzyme; receptor; PCR primer; ss.
 XX Synthetic.
 OS Bacillus subtilis.
 OS WO9822598-A1.
 PN
 XX
 XX 28-MAY-1998.
 XX
 XX 18-NOV-1997; 97WO-US21084.
 XX
 XX 12-JUN-1997; 97US-0049441.
 PR 18-NOV-1996; 96US-0749521.
 XX
 XX (NOVO) NOVO NORDISK BIOTECH INC.
 PA
 PI Adams LF, Brown S, Sloma A, Sternberg D;
 XX WPI; 1998-312483/27.
 DR
 XX Producing polypeptides in Bacillus cells - which are modified to
 PT reduce production of surfactin, useful for, e.g. producing hormones
 PT or enzymes
 XX Example 5; Page 25; 43pp; English.
 PS
 XX PCR primers AAV38593-94 and AAV38595-96 were used to amplify the
 CC different portions of the amyE gene of Bacillus subtilis strain Al64.
 CC The portions were cloned to create a partially deleted amyE gene. This
 CC was used in the course of the invention. The specification describes the
 CC production of a polypeptide. This comprises culturing a mutant Bacillus
 CC that includes nucleic acid encoding the polypeptide and nucleic acid
 CC containing a modified form of at least 1 gene involved in biosynthesis
 CC of a surfactin or its isoforms. The mutant organism produces less
 CC surfactin, or isoform, than wild-type Bacillus. The method may be used
 CC to produce hormones, enzymes and receptors, which are either
 CC heterologous or homologous.
 XX
 XX Sequence 34 BP; 5 A; 10 C; 11 G; 8 T; 0 other;
 SQ

Query Match 100.0%; Score 11; DB 19; Length 34;
 Best Local Similarity 100.0%; Pred. No. 3.4e+02;
 Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GCCTTAAGGCC 11
 Db 3 gccttaaggcc 13

RESULT 12
 AAV82401
 ID AAV82401 standard; DNA; 34 BP.
 XX
 AC AAV82401;
 XX
 XX 12-APR-1999 (first entry)
 DF
 XX Bacillus subtilis alpha-amylase amyE gene PCR primer 19.
 DE
 XX Alpha-amylase; amyE gene; protease; PCR; primer; ss.
 KW
 XX Synthetic.
 OS Bacillus subtilis.
 OS WO9856927-A2.
 PN
 XX
 XX 17-DEC-1998.
 PD
 XX 09-JUN-1998; 98WO-US12005.
 XX
 XX

PR 12-JUN-1997; 97US-0873479.
 XX (NOVO) NOVO NORDISK BIOTECH INC.
 PA
 XX Christianson L, Sloma A;
 PI
 XX WPI; 1999-080908/07.
 DR
 XX Novel protease from Bacillus subtilis LC20 - useful in laundry and
 PT dishwashing detergents and for leather processing
 PT
 XX Example 5; Page 30; 77pp; English.
 PS
 XX Primers 19 and 20 (see AAV82401-02) are designed for use in the PCR
 CC amplification of a downstream portion (nucleotides 445-953,
 CC where the A of the ATG start codon is +1) of the alpha-amylase
 CC amyE gene from Bacillus subtilis Al64 chromosomal DNA. Primer 19
 CC includes SfiI and NotI sites. An upstream portion of the amyE gene
 CC (nucleotides -421 to +77) was amplified using primers 17 and 18
 CC (see AAV82399-400). The 2 fragments were then spliced by PCR using
 CC primers 17 and 20. A splicing by overlap extension (SOE) reaction
 CC was performed to create a deleted version of the amyE gene. Mutant
 CC host cells having a deleted spoIIAC gene, a deleted amyE gene and
 CC deleted protease genes have been constructed. These can be used in
 CC the production of heterologous polypeptides. The invention also
 CC relates to Bacillus proteases (see AAV89547-48) suitable for use in
 CC detergent compositions and in leather processing.
 XX
 XX Sequence 34 BP; 5 A; 10 C; 11 G; 8 T; 0 other;
 SQ

Query Match 100.0%; Score 11; DB 20; Length 34;
 Best Local Similarity 100.0%; Pred. No. 3.4e+02;
 Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GCCTTAAGGCC 11
 Db 3 gccttaaggcc 13

RESULT 13
 AAV38594/C
 ID AAV38594 standard; DNA; 36 BP.
 XX
 AC AAV38594;
 XX
 XX 17-SEP-1998 (first entry)
 DT
 XX PCR primer 18 used to amplify a portion of the amyE gene.
 DE
 XX amyE gene; strain Al64; mutant Bacillus; biosynthesis; surfactin;
 KW production; hormone; enzyme; receptor; PCR primer; ss.
 XX
 XX Synthetic.
 OS Bacillus subtilis.
 OS WO9822598-A1.
 PN
 XX 28-MAY-1998.
 XX
 XX 18-NOV-1997; 97WO-US21084.
 XX
 XX 12-JUN-1997; 97US-0049441.
 PR 18-NOV-1996; 96US-0749521.
 XX
 XX (NOVO) NOVO NORDISK BIOTECH INC.
 PA
 XX Adams LF, Brown S, Sloma A, Sternberg D;
 PI
 XX WPI; 1998-312483/27.
 DR
 XX Producing polypeptides in Bacillus cells - which are modified to
 PT reduce production of surfactin, useful for, e.g. producing hormones
 PT

PT or enzymes
 XX Example 5; Page 25; 43pp; English.
 XX
 CC PCR primers AAV38593-94 and AAV38595-96 were used to amplify the
 CC different portions of the amyE gene of *Bacillus subtilis* strain Al64.
 CC The portions were cloned to create a partially deleted amyE gene. This
 CC was used in the course of the invention. The specification describes the
 CC production of a polypeptide. This comprises culturing a mutant *Bacillus*
 CC that includes nucleic acid encoding the polypeptide and nucleic acid
 CC containing a modified form of at least 1 gene involved in biosynthesis
 CC of a surfactin or its isoforms. The mutant organism produces less
 CC surfactin, or isoform, than wild-type *Bacillus*. The method may be used
 CC to produce hormones, enzymes and receptors, which are either
 CC heterologous or homologous.
 XX
 SQ Sequence 36 BP; 11 A; 12 C; 10 G; 3 T; 0 other;

Query Match 100.0%; Score 11; DB 19; Length 36;
 Best Local Similarity 100.0%; Pred. No. 3.4e+02;
 Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GCCTTAAGGCG 11
 |||||
 Db 22 GCCTTAAGGCG 12

RESULT 14
 AAV82400/c
 ID AAV82400 standard; DNA; 36 BP.
 XX
 AC AAV82400;
 XX
 DT 12-APR-1999 (first entry)
 XX
 DE *Bacillus subtilis* alpha-amylase amyE gene PCR primer 19.
 XX
 KW Alpha-amylase; amyE gene; protease; PCR; primer; ss.
 XX
 OS Synthetic.
 OS *Bacillus subtilis*.
 XX
 PN W09856927-A2.
 XX
 PD 17-DEC-1998.
 XX
 PF 09-JUN-1998; 98WO-US12005.
 XX
 PR 12-JUN-1997; 97US-0873479.
 XX
 PA (NOVO) NOVO NORDISK BIOTECH INC.
 XX
 PI Christianson L, Sloma A;
 XX
 DR WPI; 1999-080908/07.
 XX
 PT Novel protease from *Bacillus subtilis* LC20 - useful in laundry and
 PT dishwashing detergents and for leather processing
 XX
 PS Example 5; Page 29; 77pp; English.
 XX
 CC Primers 17 and 18 (see AAV82399-400) are designed for use in the PCR
 CC amplification of an upstream portion (nucleotides -421 to +77,
 CC where the A of the ATG start codon is +1) of the alpha-amylase
 CC amyE gene from *Bacillus subtilis* Al64 chromosomal DNA. Primer 18
 CC (nucleotides 445-953) was amplified using primers 19 and 20 (see
 CC AAV82401-02). The 2 fragments were then spliced by PCR using
 CC primers 17 and 20. A splicing by overlap extension (SOE) reaction
 CC was performed to create a deleted version of the amyE gene. Mutant
 CC host cells having a deleted *spoIIAC* gene, a deleted amyE gene and
 CC deleted protease genes have been constructed. These can be used in

CC the production of heterologous polypeptides. The invention also
 CC relates to *Bacillus* proteases (see AAV89547-48) suitable for use in
 CC detergent compositions and in leather processing.
 XX
 SQ Sequence 36 BP; 11 A; 12 C; 10 G; 3 T; 0 other;

Query Match 100.0%; Score 11; DB 20; Length 36;
 Best Local Similarity 100.0%; Pred. No. 3.4e+02;
 Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GCCTTAAGGCG 11
 |||||
 Db 22 GCCTTAAGGCG 12

RESULT 15
 AA223328
 ID AA223328 standard; DNA; 37 BP.
 XX
 AC AA223328;
 XX
 DT 06-DEC-1999 (first entry)
 XX
 DE *Bacillus licheniformis* amyL gene promoter PCR primer 1.
 XX
 KW Tandem promoter; consensus promoter; enzyme production; hormone; amyL;
 KW antibody; reporter; marker gene; cell regulation; serine protease;
 KW SAVINASE; PCR primer; ss.
 XX
 OS Synthetic.
 OS *Bacillus licheniformis*.
 XX
 PN W09943835-A2.
 XX
 PD 02-SEP-1999.
 XX
 PF 26-FEB-1999; 99WO-US04360.
 XX
 PR 26-FEB-1998; 98US-0031442.
 XX
 PA (NOVO) NOVO NORDISK BIOTECH INC.
 XX
 PI Widner W, Sloma A, Thomas MD;
 XX
 DR WPI; 1999-561370/47.
 XX
 PT Production of polypeptide in *Bacillus* using specific promoters,
 PT particularly for producing enzymes -
 XX
 PS Example 6; Page 26; 90pp; English.
 XX
 CC This invention describes a novel method for the production of a
 CC polypeptide in *Bacillus* using specific tandem or consensus promoters.
 CC The method is used to produce homologous or particularly heterologous
 CC proteins, particularly enzymes (specifically serine protease, maltogenic
 CC alpha-amylase and pullulanase), but also hormones, antibodies, reporters
 CC etc. The specified promoters provide increased expression of the sequence
 CC which encodes the polypeptide of the invention. After incorporation of
 CC the nucleic acid construct of the invention, any marker gene may be
 CC deleted, resulting in a cell that is preferred for environmental and
 CC regulatory regions. This sequence represents a PCR primer used to amplify
 CC the *Bacillus licheniformis* alpha-amylase amyL promoter which is used in
 CC the construction of an amyL promoter/serine protease, SAVINASE (RTM) gene
 CC expression cassette described in the method of the invention.
 XX
 SQ Sequence 37 BP; 4 A; 12 C; 10 G; 11 T; 0 other;

Query Match 100.0%; Score 11; DB 20; Length 37;
 Best Local Similarity 100.0%; Pred. No. 3.4e+02;
 Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GCCTTAAGGC 11
| | | | | | | | | |
Db 5 gccttaaggc 15

Search completed: January 17, 2002, 12:02:19
Job time: 17726 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: January 17, 2002, 08:54:00 ; Search time 6501.33 Seconds
(without alignments)
18.181 Million cell updates/sec

Title: US-09-242-202a-10_COPY_1_11

Perfect score: 11

Sequence: 1 GCCTTAAGGC 11

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 11351937 seqs, 5372889281 residues

Total number of hits satisfying chosen parameters: 22703874

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

EST:*
1: em_estfun:*
2: em_esthum:*
3: em_estin:*
4: em_estom:*
5: em_estpl:*
6: em_estba:*
7: em_estro:*
8: em_estov:*
9: em_btc:*
10: gb_estl:*
11: gb_est2:*
12: gb_htc:*
13: gb_gss:*
14: em_gss_fun:*
15: em_gss_hum:*
16: em_gss_inv:*
17: em_gss_pln:*
18: em_gss_pro:*
19: em_gss_rod:*
20: em_gss_vrt:*
21: em_gss_other:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	11	100.0	100	AA093528	AA093528 c10250.se
2	11	100.0	115	AZ119770	AZ119770 RPCI-23-4
3	11	100.0	130	AQ355689	AQ355689 CITBI-E1-
4	11	100.0	136	AA345723	AA345723 EST51785
5	11	100.0	140	AZ123745	AZ123745 RPCI-23-4
6	11	100.0	147	N27010	N27010 yw73a09.s1
7	11	100.0	149	BE245846	BE245846 TCBAP1E19
8	11	100.0	158	N41203	N41203 HRI179-F.Adu
9	11	100.0	159	BE244007	BE244007 TCBAP1E04
10	11	100.0	165	AQ012949	AQ012949 CIT-HSP-2
11	11	100.0	172	AW308236	AW308236 3366 MARC
12	11	100.0	181	AI303597	AI303597 u181a12.x

c 13	11	100.0	182	10	AI800916
c 14	11	100.0	185	11	BF605530
c 15	11	100.0	199	13	AZ075637
c 16	11	100.0	203	10	AA303217
c 17	11	100.0	207	10	AA498912
c 18	11	100.0	208	10	BE698384
c 19	11	100.0	213	11	F18421
c 20	11	100.0	214	10	AW197056
c 21	11	100.0	223	10	BB143920
c 22	11	100.0	227	10	AV321312
c 23	11	100.0	230	11	F25526
c 24	11	100.0	231	10	AW920467
c 25	11	100.0	231	11	N23615
c 26	11	100.0	232	11	BG552774
c 27	11	100.0	232	11	BI262529
c 28	11	100.0	234	10	BB152224
c 29	11	100.0	235	10	AI306150
c 30	11	100.0	241	10	AV062684
c 31	11	100.0	241	10	AV146659
c 32	11	100.0	243	11	Z38396
c 33	11	100.0	246	10	AA094468
c 34	11	100.0	251	10	AU127469
c 35	11	100.0	251	13	AZ366083
c 36	11	100.0	252	10	AV649344
c 37	11	100.0	252	10	BB606444
c 38	11	100.0	254	10	AA576933
c 39	11	100.0	255	11	N83262
c 40	11	100.0	255	11	BF60539
c 41	11	100.0	256	10	AV281033
c 42	11	100.0	257	11	BF509446
c 43	11	100.0	259	10	AI350354
c 44	11	100.0	259	10	BB390531
c 45	11	100.0	259	13	AZ217098

ALIGNMENTS

RESULT 1

AA093528	AA093528	100 bp	mRNA	EST	25-OCT-1996
LOCUS	c10250.seq.F Human fetal heart, Lambda ZAP Express Homo sapiens				
DEFINITION	CDNA 5', mRNA sequence.				
ACCESSION	AA093528				
VERSION	AA093528.1	GI:1639161			
KEYWORDS	EST.				
SOURCE	human.				
ORGANISM	Homo sapiens				
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.				
AUTHORS	Liew,C.C.				
TITLE	CDNAs from fetal heart (1996)				
JOURNAL	Unpublished (1996)				
COMMENT	Contact: Liew CC Department of Laboratory Medicine and Pathobiology University of Toronto Banting Institute, 100 College St., Toronto, Ontario, M5G1L5 Tel: 4169788758 Fax: 4169785650 Email: liewccc@utcc.utoronto.ca				
PCR Primers					
FORWARD:	5' GCCAAGCTCGAATTAACCCCTCACTAAGG 3'				
BACKWARD:	5' CCAGTGAATTGTAATACGACTCACTATAGGCG 3'				
Seq primer:	5' GAAATTAACCTCACTAAGG 3'				

FEATURES

1..100
Location/Qualifiers
1..100
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone_lib="Human fetal heart, Lambda ZAP Express"
/lab_host="E. coli XL1-Blue"
/note="Vector: Lambda ZAP Express; Site:1: EcoRI; Site_2: XhoI; mRNA was purified from human fetal hearts (8-10

weeks). cDNA was synthesized using a XhoI-Oligo dt adaptor-primer. EcoRI adaptors were ligated, followed by digestion with XhoI, for directional cloning into predigested lambda ZAP Express."

BASE COUNT

10 a 39 c 27 g 24 t

Query Match

Best Local Similarity 100.0%; Score 11; DB 10; Length 100;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GCCTTAAGGC 11

Db 78 GCCTTAAGGC 88

RESULT 2

AZ119770/c
LOCUS AZ119770 115 bp DNA 12-MAY-2000
DEFINITION RPCI-23-479A23-TV RPCI-23 Mus musculus genomic clone RPCI-23-479A23
ACCESSION AZ119770
VERSION AZ119770.1 GI:7785028
KEYWORDS GSS.
SOURCE house mouse.
ORGANISM Mus musculus

REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
AUTHORS Zhao, S., Nierman, W., Feldblyum, T., Malek, J., Shatsman, S., Akinret
B., Levins, M., McGann, S., Tsegaye, G., Geer, K., Kroi, M., de Jong, P.
and Fraser, C.M.
TITLE Mouse BAC End Sequences from Library RPCI-23
JOURNAL Unpublished (1999)
COMMENT Other_GSSs: RPCI-23-479A23.TJ
Contact: Shaying Zhao
Department of Eukaryotic Genomics
The Institute for Genomic Research
9712 Medical Center Dr., Rockville, MD 20850, USA
Tel: 301 838 0200
Fax: 301 838 0208
Email: szhao@tigr.org

Clones are derived from the mouse BAC library RPCI-23. For BAC library availability, please contact Pieter de Jong (pieter@dejong.med.buffalo.edu). Clones may be purchased from BACPAC Resources (http://bacpac.med.buffalo.edu/orderingframe.htm) or from Resea ch Genetics (info@resgen.com). BAC end page: http://www.tigr.org/tdb/bac_ends/mouse/bac_end_intro.html
Plate: 479 row: A column: 23
Seq primer: T7
Class: BAC ends.

FEATURES

source
1. .115
Location/Qualifiers
/organism="Mus musculus"
/strain="C57BL/6J"
/db_xref="taxon:10090"
/clone="RPCI-23-479A23"
/clone_lib="RPCI-23"
/sex="Female"
/lab_host="DH10B"
/note="Organ: Kidney/Brain; Vector: pBACe3.6; Site_1: EcoRI; Site_2: EcoRI; Female C57BL/6J mouse kidney and/or brain genomic DNA was isolated and partially digested with a combination of EcoRI and EcoRI Methylase. Size selected DNA was cloned into the pBACe3.6 vector at the EcoRI sites. The ligation products were transformed into DH10B electrocompetent cells (BRL Life Technologies)."

BASE COUNT

35 a 32 c 28 g 20 t

Query Match

100.0%; Score 11; DB 13; Length 115;

Best Local Similarity 100.0%; Pred. No. 7.5e+03;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GCCTTAAGGC 11

Db 114 GCCTTAAGGC 104

RESULT 3

AQ355689
LOCUS CITBI-E1-2538F8.TR CITBI-E1 Homo sapiens genomic clone 2538F8, DNA
DEFINITION sequence.
ACCESSION AQ355689
VERSION AQ355689.1 GI:4182862
KEYWORDS GSS.
SOURCE human.
ORGANISM Homo sapiens

REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
AUTHORS Zhao, S., Adams, M.D., Nierman, W., Malek, J., Shizuya, H., Simon, M. and
Venter, J.C.
TITLE Use of BAC End Sequences from CalTech Libraries for Sequence-Ready
Map Building
JOURNAL Unpublished (1997)
COMMENT Contact: Shaying Zhao, William Nierman, Mark Adams
Department of Eukaryotic Genomics
The Institute for Genomic Research
9712 Medical Center Dr., Rockville, MD 20850
Tel: 301 838 0200
Fax: 301 838 0208
Email: hbeet@tigr.org

Clones are available from Research Genetics (info@resgen.com). BAC end search page:
http://www.tigr.org/tdb/humgen/bac_end_search/bac_end_search.html.
Seq primer: M13 Reverse
Class: BAC ends.

FEATURES

source
1. .130
Location/Qualifiers
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="2538F8"
/clone_lib="CITBI-E1"
/sex="male"
/cell_type="sperm"
/note="Vector: pBelOBAC11; Site_1: EcoRI; Site_2: EcoRI;
Caltech Human BAC Library D"

BASE COUNT 42 a 23 c 21 g 44 t

Query Match

Best Local Similarity 100.0%; Score 11; DB 13; Length 130;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GCCTTAAGGC 11

Db 23 GCCTTAAGGC 33

RESULT 4

AA345723
LOCUS EST51785 Gall bladder I Homo sapiens cDNA 5', end similar to similar
DEFINITION to prollylcarboxypeptidase, mRNA sequence.
ACCESSION AA345723
VERSION AA345723.1 GI:1997971
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens

REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE 1 (bases 1 to 136)
 AUTHORS Adams,M.D., Kerlavage,A.R., Fleischmann,R.D., Fuldner,R.A., Bult
 C.J., Lee,N.H., Kirkness,E.F., Weinstock,K.G., Gocayne,J.D., White
 O., Sutton,G., Blake,J.A., Brandon,R.C., Man-Wai,C., Clayton,R.A.,
 Cline,T.R., Cotton,M.D., Earle-Hughes,J., Fine,L.D., Fitzgerald
 J.M., Fitchugh,W.M., Fritchman,J.L., Geoghagen,N.S., Glodek,A.,
 Gnehm,C.L., Hanna,M.C., Hedblom,E., Hinkle,P.S., Jr., Kelley,J.M.,
 Kelley,J.C., Liu,L.-I., Marmaros,S.M., Merrick,J.M.,
 Moreno-Palancas,R.F., McDonald,L.A., Nguyen,D.T., Pelligrino,S.M.,
 Phillips,C.A., Ryder,S.E., Scott,J.L., Saudek,D.M., Shirley,R.,
 Small,K.V., Spriggs,T.A., Utterback,T.R., Weidman,J.F., Li,Y.,
 Bednarik,D.P., Cao,L., Cepeda,M.A., Coleman,T.A., Collins,E.J.,
 Dimke,D., Feng,D.-F., Ferrie,A., Fischer,C., Hastings,G.A., He,W.W.,
 Hu,J.S., Greene,J.M., Gruber,J., Hudson,P., Kim,A.K., Kozak,D.L.,
 Kunsch,C., Hungjun,J., Li,H., Weissner,P.S., Olsen,H., Raymond,L.,
 Wei,Y.F., Wing,J., Xu,C., Yu,G.L., Ruben,S.M., Dillion,P.J., Pannon
 M.R., Rosen,C.A., Haseltine,W.A., Fields,C., Fraser,C.M. and
 Venter,J.C.
 TITLE Initial assessment of human gene diversity and expression patterns
 based upon 83 million nucleotides of cDNA sequence
 JOURNAL Nature 377 (6547 Suppl), 3-174 (1995)
 MEDLINE 96026280
 COMMENT Contact: Kerlavage, AR
 Bioinformatics
 The Institute for Genomic Research
 9712 Medical Center Drive, Rockville, MD 20850 USA
 Tel: 3018699056
 Fax: 3018699423
 Email: arkerlav@tigr.org
 For clone availability, additional sequence and expression
 information related to this EST, please check the TIGR Human Gene
 Index (<http://www.tigr.org/tdb/hgi/hgi.html>)
 Seq primer: M13 Reverse

FEATURES
 source
 1. .136
 /organism="Homo sapiens"
 /db_xref="ATCC (inhost):147260"
 /db_xref="taxon:9606"
 /clone_lib="Gall bladder I"
 /sex="female"
 /dev_stage="adult, 25 yrs"
 /note="Organ: gall bladder; Vector: pBluescript SK-;
 Site_1: EcoRI; Site_2: XhoI"
 BASE COUNT 18 a 59 c 30 g 25 t 4 others
 ORIGIN

Query Match 100.0%; Score 11; DB 10; Length 136;
 Best Local Similarity 100.0%; Pred. No. 7.5e+03;
 Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GCCTTAAGGCG 11
 |||||
 Db 89 GCCTTAAGGCG 99

RESULT 5
 AZ123745/c 140 bp DNA GSS 12-MAY-2000
 LOCUS
 DEFINITION RPCI-23-479M23-TV RPCI-23 Mus musculus genomic clone RPCI-23-479M23
 , DNA sequence.
 ACCESSION AZ123745
 VERSION AZ123745.1 GI:7792734
 KEYWORDS GSS.
 SOURCE house mouse.
 ORGANISM Mus musculus
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 1 (bases 1 to 140)
 REFERENCE
 AUTHORS Zhao,S., Nierman,W., Feldblyum,T., Malek,J., Shatsman,S., Akinret
 ,B., Levins,M., McGann,S., Tsegaye,G., Geer,K., Krol,M., de Jong,P.
 and Fraser,C.M.
 TITLE Mouse BAC End Sequences from Library RPCI-23

JOURNAL COMMENT
 Unpublished (1999)
 Contact: Shaying Zhao
 Department of Eukaryotic Genomics
 The Institute for Genomic Research
 9712 Medical Center Dr., Rockville, MD 20850, USA
 Tel: 301 838 0200
 Fax: 301 838 0208
 Email: szhao@tigr.org
 Clones are derived from the mouse BAC library RPCI-23. For BAC
 library availability, please contact Pieter de Jong
 (pieterdejong.med.buffalo.edu). Clones may be purchased from
 BACPAC Resources (<http://bacpac.med.buffalo.edu/orderingframe.htm>)
 or from Resea ch Genetics (info@resgen.com). BAC end page:
http://www.tigr.org/tdb/bacends/mouse/bac_end_intro.html
 Plate: 479 row: M column: 23
 Seq primer: T7
 Class: BAC ends.
 Location/Qualifiers
 1. .140
 /organism="Mus musculus"
 /strain="C57BL/6J"
 /db_xref="taxon:10090"
 /clone="RPCI-23-479M23"
 /clone_lib="RPCI-23"
 /sex="Female"
 /lab_host="DH10B"
 /note="Organ: Kidney/Brain; Vector: pBACE3.6; Site_1:
 EcoRI; Site_2: EcoRI; Female C57BL/6J mouse kidney and/or
 brain genomic DNA was isolated and partially digested
 with a combination of EcoRI and EcoRI Methylase. Size
 selected DNA was cloned into the pBACE3.6 vector at the
 EcoRI sites. The ligation products were transformed into
 DH10B electrocompetent cells (BRL Life Technologies)."
 BASE COUNT 45 a 39 c 34 g 22 t
 ORIGIN

Query Match 100.0%; Score 11; DB 13; Length 140;
 Best Local Similarity 100.0%; Pred. No. 7.5e+03;
 Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GCCTTAAGGCG 11
 |||||
 Db 139 GCCTTAAGGCG 129

RESULT 6
 N27010 147 bp mRNA EST 29-DEC-1995
 LOCUS yw73a09.s1 Soares.placenta.8to9weeks_2NbHP8to9W Homo sapiens CDNA
 DEFINITION Clone IMAGE:257848 3', similar to gb:M30142 COMPLEMENT
 DECAT-ACCELERATING FACTOR 1 PRECURSOR (HUMAN);, mRNA sequence.
 N27010
 VERSION N27010.1 GI:1141358
 KEYWORDS EST.
 SOURCE human.
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 1 (bases 1 to 147)
 REFERENCE
 AUTHORS Hillier,L., Clark,N., Dubuque,T., Elliston,K., Hawkins,M., Holman
 M., Hultman,M., Kucaba,T., Le,M., Lennon,G., Marra,M., Parsons,J.,
 Rifkin,L., Rohlfing,T., Soares,M., Tan,F., Trevaskis,E., Waterston
 R., Williamson,A., Wohldmann,P. and Wilson,R.
 The WashU-Merck EST Project
 Unpublished (1995)
 TITLE
 JOURNAL
 COMMENT Contact: Willson RK
 Washington University School of Medicine
 4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
 Tel: 314 286 1800
 Fax: 314 286 1810
 Email: est@watson.wustl.edu
 High quality sequence starts: 1

High quality sequence stops: 1
Source: IMAGE Consortium, LLNL
This clone is available royalty-free through LLNL; contact the
IMAGE Consortium (info@image.llnl.gov) for further information.
Trace Consortium (info@image.llnl.gov) for further information.
Possible reversed clone: polyt not found
Seq primer: m13 -40 forward
High quality sequence stop: 1.
Location/Qualifiers
1. .147
/organism="Homo sapiens"
/db_xref="GDB:3887458"
/db_xref="taxon:9606"
/clone="IMAGE:257848"
/clone_lib="Soares_placenta_8to9weeks_2NBP8to9W"
/dev_stage="two placentae: one from 8 weeks and another
from 9 weeks post conception"
/lab_host="DH10B (ampicillin resistant)"
/note="Organ: placenta; Vector: pT7T3D (Pharmacia) with a
modified polylinker; Site_1: Not I; Site_2: Eco RI; 1st
strand cDNA was primed with a Not I - Oligo(dT) primer [5'
TGTACCAATCTGAAGTGGAGCGCGCGATTTTTTTTTTTT 3'],
double-stranded cDNA was size selected, ligated to Eco RI
adapters (Pharmacia), digested with Not I and cloned into
the Not I and Eco RI sites of a modified pT7T3 vector
(Pharmacia). Library constructed by Bento Soares and
M.Patima Bonaldo."
BASE COUNT 41 a 26 c 33 g 36 t 11 others
ORIGIN

Query Match 100.0%; Score 11; DB 11; Length 147;
Best Local Similarity 100.0%; Pred. No. 7.5e+03;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 GCCTTAAGGCG 11
|||||
Db 102 GCCTTAAGGCG 112

RESULT 7
BE245846/c
LOCUS
DEFINITION
TCBAP1E1938 Pediatric pre-B cell acute lymphoblastic leukemia
Baylor-HGSC project-TCBA Homo sapiens cDNA clone TCBAP1938, mRNA
sequence.
ACCESSION BE245846
VERSION BE245846.1 GI:9097594
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 149)
AUTHORS Wei,Y., Tsang,Y.T.M., Mei,G., Ku,J.M., Ali-Osman Jr.,F.R., Muzny,D.,
Bouck,J., Gibbs,R.A. and Margolin,J.F.
TITLE Pediatric Leukemia cDNA Sequencing Project
JOURNAL Unpublished (2000)
COMMENT Contact: Dr. Judith F. Margolin
Texas Children's Cancer Center and Human Genome Sequencing Center
at Baylor College of Medicine
1102 Bates, MC3-3320 Houston, TX 77030, USA
Tel: 832-824-4536
Fax: 832-825-4038
Email: clones@xccc.org
Seq primer: M13 primer
Location/Qualifiers
1. .149
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="TCBAP1938"
/clone_lib="Pediatric pre-B cell acute lymphoblastic
leukemia Baylor-HGSC project=TCBA"

/sex="male"
/tissue_type="leukopheresis"
/cell_type="pre-B cell"
/dev_stage="pediatric 2 years"
/lab_host="DH10B"
/note="Vector: lambda pSB; Site_1: BamHI; Site_2: EcoRI;
First strand cDNA was primed with an anchored
XhoI-oligo(dT) primer [5'GGAGGACTCGAGCGCGCGAGGAG(T)VN
3'; V=A,C,G; N=A,C,G,T] and then dg tailed. Second
strand was primed with a BamHI-dc primer
[5'AGAGAGCTCGATCCGCGCGCAATAATAAT(C) 3'].
Double-stranded cDNA was then digested with BamHI and
XhoI and directionally cloned into the BamHI and SalI
sites of lambda pSB vector. Library went through one
round of normalization. Library was constructed by Wei
Yu."
BASE COUNT 35 a 56 c 46 g 10 t 2 others
ORIGIN

Query Match 100.0%; Score 11; DB 10; Length 149;
Best Local Similarity 100.0%; Pred. No. 7.5e+03;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 GCCTTAAGGCG 11
|||||
Db 50 GCCTTAAGGCG 40

RESULT 8
N41203/c
LOCUS
DEFINITION
HR179-F Adult heart, Clontech Homo sapiens cDNA clone R179-F, mRNA
sequence.
ACCESSION N41203
VERSION N41203.1 GI:1164801
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 158)
AUTHORS Wayne,M.M.Y., Cheung,H.K.Y., Lam,W.Y., Law,P.T.W., Lo,A.S.Y., Lui
,V.W.Y., Luk,S.C.W., Tsui,S.K.W., Tung,C.K.C., Yan,N.Y.H., Liew
,C.C. and Lee,C.Y.
TITLE Gene expression of adult human heart as revealed by random
sequencing of cDNA library
JOURNAL Miami Winter Biotechnol. Symp. Proc. 6, 90 (1995)
COMMENT Contact: Wayne Mary M.Y.
Department of Biochemistry
The Chinese University of Hong Kong
Rm 302C, Basic Medical Science Building, The Chinese University of
Hong Kong, Shatin, N.T., Hong Kong.
Tel: 8526096874
Fax: 8526035123
Email: bl33723@vax.csc.cuhk.hk
Seq primer: GTGGCGGAGCTCTCTGGAGCC.
Location/Qualifiers
1. .158
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="R179-F"
/clone_lib="Adult heart, Clontech"
/lab_host="E. coli Y1090"
/note="Vector: Lambda gtl1; Site_1: EcoRI; Site_2: EcoRI"

BASE COUNT 49 a 50 c 35 g 24 t
ORIGIN

Query Match 100.0%; Score 11; DB 11; Length 158;
Best Local Similarity 100.0%; Pred. No. 7.5e+03;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;


```

Qy 1 GCCTTAAGGCG 11
Db 141 GCCTTAAGGCG 131

RESULT 9
BE244007/c
LOCUS
DEFINITION
  BE244007 159 bp mRNA EST 15-NOV-2000
  TCBAP1E0434 Pediatric pre-B cell acute lymphoblastic leukemia
  Baylor-HGSC project-TCBA Homo sapiens cDNA clone TCBAP0434, mRNA
  sequence.
ACCESSION
  BE244007
VERSION
  BE244007.1 GI:9095747
SOURCE
  EST.
ORGANISM
  Homo sapiens
  Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
  Mammalia; Eutheria; Primates; Catarrhini; Hominoidea; Homo.
REFERENCE
  1 (bases 1 to 159)
  Wei, Y., Tsang, Y.T.M., Mei, G., Ku, J.M., Ali-Osman Jr., F.R., Muzny, D.,
  Bouck, J., Gibbs, R.A. and Margolin, J.F.
  Pediatric Leukemia cDNA Sequencing Project
  Unpublished (2000)
  Contact: Dr. Judith F. Margolin
  Texas Children's Cancer Center and Human Genome Sequencing Center
  at Baylor College of Medicine
  1102 Bates, MC3-3320 Houston, TX 77030, USA
  Tel: 832-824-4536
  Fax: 832-825-4038
  Email: clones@ccc.org
  Seq primer: M13 primer.
  Location/Qualifiers
    1..159
      /organism="Homo sapiens"
      /db_xref="taxon:9606"
      /clone_lib="TCBAP0434"
      /clone_lib="Pediatric pre-B cell acute lymphoblastic
      leukemia Baylor-HGSC project-TCBA"
      /sex="male"
      /tissue_type="leukopheresis"
      /cell_type="pre-B cell"
      /dev_stage="pediatric 2 years"
      /lab_host="DH108"
      /notes="Vector: lambda pSB; Site_1: BamHI; Site_2: EcoRI;
      First strand cDNA was primed with an anchored
      XhoI-oligo(dT) primer [5'GGAGGACTCGAGCGCGCGAGGAG(T)VN
      3'; V=A,C,G; N=A,C,G,T] and then dG tailed. Second
      strand was primed with a BamHI-dC primer
      [5'AGAGCTCGGATCCGCGCGCAATAATAAT(C) 3'].
      Double-stranded cDNA was then digested with BamHI and
      XhoI and directionally cloned into the BamHI and SalI
      sites of lambda pSB vector. Library went through one
      round of normalization. Library was constructed by Wei
      Yu"
BASE COUNT 48 a 48 c 47 g 12 t 4 others
ORIGIN

Query Match 100.0%; Score 11; DB 10; Length 159;
Best Local Similarity 100.0%; Pred. No. 7.5e+03;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GCCTTAAGGCG 11
Db 42 GCCTTAAGGCG 32

RESULT 10
AQ012949
LOCUS
DEFINITION
  AQ012949 165 bp DNA GSS 06-JUN-1998
  CIT-HSP-2300F24.TF CIT-HSP Homo sapiens genomic clone 2300F24, DNA
  sequence.
ACCESSION
  AQ012949

Qy 1 GCCTTAAGGCG 11
Db 42 GCCTTAAGGCG 32

RESULT 11
AW308236/c
LOCUS
DEFINITION
  AW308236 172 bp mRNA EST 25-APR-2001
  3366 MARC 1BOV Bos taurus cDNA 5', mRNA sequence.
ACCESSION
  AW308236
VERSION
  AW308236.1 GI:6720599
KEYWORDS
  EST.
SOURCE
  COW.
  Bos taurus
  Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
  Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
  Bovinae; Bos.
  1 (bases 1 to 172)
  Smith, P.L., Grosse, W.M., Freking, B.A., Roberts, A.J., Stone, R.T.,
  Casas, E., Wray, J.E., White, J., Cho, J., Fahrenkrug, S.C., Bennett
  , G.L., Heaton, M.P., Laegreid, W.W., Rohrer, G.A., Chitko-McKown, C.G.,
  Perle, G., Holt, I., Karamycheva, S., Liang, F., Quackenbush, J. and
  Keele, J.W.
  Sequence evaluation of four pooled-tissue normalized bovine cDNA
  libraries and construction of a gene index for cattle
  Genome Res. 11 (4), 626-630 (2001)
  21180013
  Contact: Smith TPL

Qy 1 GCCTTAAGGCG 11
Db 60 GCCTTAAGGCG 70

Query Match 100.0%; Score 11; DB 13; Length 165;
Best Local Similarity 100.0%; Pred. No. 7.5e+03;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GCCTTAAGGCG 11
Db 60 GCCTTAAGGCG 70

RESULT 11
AW308236/c
LOCUS
DEFINITION
  AW308236 172 bp mRNA EST 25-APR-2001
  3366 MARC 1BOV Bos taurus cDNA 5', mRNA sequence.
ACCESSION
  AW308236
VERSION
  AW308236.1 GI:6720599
KEYWORDS
  EST.
SOURCE
  COW.
  Bos taurus
  Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
  Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
  Bovinae; Bos.
  1 (bases 1 to 172)
  Smith, P.L., Grosse, W.M., Freking, B.A., Roberts, A.J., Stone, R.T.,
  Casas, E., Wray, J.E., White, J., Cho, J., Fahrenkrug, S.C., Bennett
  , G.L., Heaton, M.P., Laegreid, W.W., Rohrer, G.A., Chitko-McKown, C.G.,
  Perle, G., Holt, I., Karamycheva, S., Liang, F., Quackenbush, J. and
  Keele, J.W.
  Sequence evaluation of four pooled-tissue normalized bovine cDNA
  libraries and construction of a gene index for cattle
  Genome Res. 11 (4), 626-630 (2001)
  21180013
  Contact: Smith TPL

```

USDA, ARS, US Meat Animal Research Center
PO Box 166, Clay Center, NE 68933-0166, USA
Tel: 402 762 4366
Fax: 402 762 4390
Email: smith@email.marc.usda.gov
Single pass sequencing. Bases called and trimmed with phred
v0.980904.e. Vector identified by cross_match with the -minscore 20
and -mismatch 12 options.
PCR Primers
FORWARD: AGGAACAGCTATGACCAT
BACKWARD: GTTTCCTCCAGTCACGAG
Plate: 147 row: J column: 11
Seq primer: ATTTAGGTGACATATAG.
Location/Qualifiers
1. .172
/organism="Bos taurus"
/db_xref="taxon:9913"
/clone_lib="MARC lBOV"
/tissue_type="pooled"
/lab_host="Dh10B"
/note="Vector: pCMV SPORT6; site_1: XbaI; site_2: XhoI;
Library made from pooled tissue from lymph node, ovary,
fat, hypothalamus, and pituitary."
40 a 61 c 47 g 24 t

BASE COUNT
ORIGIN

Query Match 100.0%; Score 11; DB 10; Length 172;
Best Local Similarity 100.0%; Pred. No. 7.5e+03;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GCCTTAAGGCG 11
|||||

Db 22 GCCTTAAGGCG 12
|||||

RESULT 12
AI303597

LOCUS
AI303597 181 bp mRNA EST 08-DEC-1998
ui81a12.x1 Sugano mouse liver mlia Mus musculus cDNA clone
IMAGE:1888798 3' similar to TR:035533 O35533 CARBOXYLESTERASE
PRECURSOR ; mRNA sequence.

ACCESSION
AI303597
AI303597.1 GI:3987347
EST.

KEYWORDS
SOURCE
ORGANISM
Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1 (bases 1 to 181)

REFERENCE
AUTHORS
Marra,M., Hillier,L., Allen,M., Bowles,M., Dietrich,N., Dubuque,T.,
Geisel,S., Kucaba,T., Lacy,M., Le,M., Martin,J., Morris,M.,
Schellenberg,K., Steptoe,M., Tan,F., Underwood,K., Moore,B.,
Theisinger,B., Wyllie,T., Lennon,G., Soares,B., Wilson,R. and
Waterston,R.
The WashU-HMI Mouse EST Project
Unpublished (1996)

TITLE
JOURNAL
COMMENT
Contact: Marra M/Mouse EST Project
WashU-HMI Mouse EST Project
Washington University School of MedicineP
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
Tel: 314 286 1800
Fax: 314 286 1810
Email: mouseest@watson.wustl.edu
This clone is available royalty-free through LNL; contact the
IMAGE Consortium (info@image.lnl.gov) for further information.
MGI:973122

Trace considered overall poor quality
Possible reversed clone: similarity on wrong strand
Seq primer: custom primer used
High quality sequence stop: 1.
Location/Qualifiers
1. .181

FEATURES
source

	/organism="Mus musculus"
	/strain="C57BL"
	/db_xref="taxon:10090"
	/clone="IMAGE:1888798"
	/clone_lib="Sugano mouse liver mlia"
	/sex="female"
	/dev_stage="adult"
	/lab_host="DH10B"
	/note="Organ: liver; Vector: pME18S-FL3; Site_1: DraIII
	(CACTGTG); Site_2: DraIII (CACTGTG); 1st strand cDNA
	was primed with an oligo(dT) primer
	[ATGTGGCCCTTTTTTTTTTTT]; double-stranded cDNA was
	ligated to a DraIII adaptor [TGTGGCTACTGG], digested
	and cloned into distinct DraIII sites of the pME18S-FL3
	vector (5' site CACTGTG, 3' site CACTGTG). XhoI should
	be used to isolate the cDNA insert. Size selection was
	performed to exclude fragments <1.5kb. Library
	constructed by Dr. Sumio Sugano (University of Tokyo
	Institute of Medical Science). Custom primers for
	sequencing: 5' end primer CTTCCTCTCTAAAAGCTGC and 3' end
	primer CGACCTGCAGCTCGAGCAC."
BASE COUNT	38 a 50 c 47 g 46 t
ORIGIN	
Query Match	100.0%; Score 11; DB 10; Length 181;
Best Local Similarity	100.0%; Pred.No. 7.4e+03;
Matches 11; Conservative	0; Mismatches 0; Indels 0; Gaps 0;
QY 1 GCCTTAAGGCC 11	
Db 81 GCCTTAAGGCC 91	
RESULT 13	
A1800916/c	
LOCUS	A1800916 182 bp mRNA EST 19-DEC-1999
DEFINITION	wg14d07.x1 Soares.NSF.F8.9W_OT_PA_P.S1 Homo sapiens cDNA clone
	IMAGE:2365069 3' similar to gb:U10320 FRUCTOSE-1,6-BISPHOSPHATASE
	(HUMAN), mRNA sequence.
ACCESSION	A1800916
VERSION	A1800916.1 GI:5366388
KEYWORDS	EST.
SOURCE	human.
ORGANISM	Homo sapiens
	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
	Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE	1 (bases 1 to 182)
AUTHORS	NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
TITLE	National Cancer Institute, Cancer Genome Anatomy Project (CGAP).
JOURNAL	Tumor Gene Index
COMMENT	Unpublished (1997) Contact: Robert Strausberg, Ph.D. Email: cgapbs@mail.nih.gov This clone is available royalty-free through LILNL ; contact the IMAGE Consortium (info@image.llnl.gov) for further information. Trace considered overall poor quality Insert length: 1617 Std Error: 0.00 Seq primer: -40UP from Glbco High quality sequence stop: 1. Location/Qualifiers 1. .182 /organism="Homo sapiens" /db_xref="taxon:9606" /clone="IMAGE:2365069" /clone_lib="Soares_NSF_F8_9W_OT_PA_P_S1" /lab_host="DH10B" /note="Organ: pooled; Vector: pT7T3D-Pac (Pharmacia) with a modified polylinker; Site_1: Not I; Site_2: Eco RI; Equal amounts of plasmid DNA from five normalized libraries were mixed, and ss circles were made in vitro. Following HAP purification, this DNA was used as tracer in a subtractive hybridization reaction. The driver was
FEATURES	
source	

PCR-amplified cDNAs from pools of 5,000 clones made from the same 5 libraries. The pools consisted of the following libraries and cloneIDs: Soares NBHSF pool 1: 309384-310919, 323208-325895 Soares NB2HP pool 1: 145032-147335, 147720-148103, 148872-149255, 15002 - 150407, 151176-153327 Soares NB2HF8-9W pool 1: 758280-760583, 772104-774407 Soares NBHPA pool 1: 304776-306311, 320136-322823, 326280-326663 Soares NBHOT pool 1: 723720-726407, 739080-740999 Subtraction by Bento Soares and M. Fatima Bonaldo."

BASE COUNT 35 a 56 c 51 g 40 t

ORIGIN

Query Match 100.0%; Score 11; DB 10; Length 182;
Best Local Similarity 100.0%; Pred. No. 7.4e+03;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GCCTTAAGGC 11
|||||
Db 85 GCCTTAAGGC 75

RESULT 14

BF605530/c 185 bp mRNA EST 25-APR-2001
LOCUS 271955 MARC 3BOV Bos taurus cDNA 5', mRNA sequence.
DEFINITION BF605530
ACCESSION BF605530.1 GI:11705237
VERSION EST.
KEYWORDS
SOURCE COW.
ORGANISM Bos taurus

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
Bovidae; Bovinae; Bos.
1 (bases 1 to 185)

AUTHORS Smith, T.P.L., Grosse, W.M., Freking, B.A., Roberts, A.J., Stone, R.T., Casas, E., Wray, J.E., White, J., Cho, J., Fahrenkrug, S.C., Bennett, G.L., Heaton, M.P., Laegreid, W.W., Rohrer, G.A., Chitko-McKown, C.G., Pertea, G., Holt, I., Karamecheva, S., Liang, F., Quackenbush, J. and Keele, J.W.

TITLE Sequence evaluation of four pooled-tissue normalized bovine cDNA libraries and construction of a gene index for cattle

JOURNAL Genome Res. 11 (4), 626-630 (2001)
MEDLINE 21180013
COMMENT Contact: Smith TPL
USDA, ARS, US Meat Animal Research Center
PO Box 166, Clay Center, NE 68933-0166, USA
Tel: 402 762 4366
Fax: 402 762 4390
Email: smith@mail.marc.usda.gov

Single pass sequencing. Bases called and alt_trimmed with phred v0.980904.e. Vector identified by cross_match with the -minscore 18 and -minmatch 12 options.

PCR Primers

FORWARD: AGGAACACGCTATGACCAT

BACKWARD: GTTTCACGTCACGACG

Plate: 55 row: F column: 24

Seq primer: ATTTAGGTGACACTATAG.

FEATURES

source

1. .185

/organism="Bos taurus"

/db_xref="taxon:9913"

/clone_lib="MARC 3BOV"

/tissue_type="pooled"

/lab_host="DH10B"

/note="Vector: pCMV SPORT6; Site 1: XbaI; Site 2: XhoI;
Library made from pooled tissue from marrow, alveolar macrophage, ovary, fetal semitendinosus muscle, and fetal longissimus muscle."

BASE COUNT 51 a 39 c 48 g 47 t

ORIGIN

Query Match 100.0%; Score 11; DB 11; Length 185;
Best Local Similarity 100.0%; Pred. No. 7.4e+03;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GCCTTAAGGC 11
|||||
Db 137 GCCTTAAGGC 127

RESULT 15

AZ075637/c.

LOCUS

DEFINITION

RPIC-23-408L19.TV RPIC-23 Mus musculus genomic clone RPIC-23-408L19

, DNA sequence.

ACCESSION

AZ075637

VERSION

AZ075637.1 GI:7368534

KEYWORDS

GSS.

SOURCE

house mouse.

ORGANISM

Mus musculus

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

1 (bases 1 to 199)

REFERENCE

AUTHORS

Zhao, S., Niernman, W., Feldblyum, T., Malek, J., Shatsman, S., Akinret, B., Levins, M., McGann, S., Tsegaye, G., Geer, K., Krol, M., de Jong, P. and Fraser, C.M.

TITLE

Mouse BAC End Sequences from Library RPIC-23

JOURNAL

Unpublished (1999)

COMMENT

Contact: Shaying Zhao

Department of Eukaryotic Genomics

The Institute for Genomic Research

9712 Medical Center Dr., Rockville, MD 20850, USA

Tel: 301 838 0200

Fax: 301 838 0208

Email: szhao@tigr.org

Clones are derived from the mouse BAC library RPIC-23. For BAC

library availability, please contact Pieter de Jong

(pieter@dejong.med.buffalo.edu). Clones may be purchased from

BACPAC Resources (http://bacpac.med.buffalo.edu/orderingframe.htm)

or from Resea ch Genetics (info@resgen.com). BAC end page:

http://www.tigr.org/tdb/bac_ends/mouse/bac_end_intro.html

Plate: 408 row: L column: 19

Seq primer: T7

Class: BAC ends.

Location/Qualifiers

1. .199

/organism="Mus musculus"

/strain="C57BL/6J"

/db_xref="taxon:10090"

/clone="RPIC-23-408L19"

/clone_lib="RPIC-23"

/sex="Female"

/lab_host="DH10B"

/note="Organ: Kidney/Brain; Vector: pBACE3.6; Site 1:

ECORI; Site 2: EcorI; Female C57BL/6J mouse kidney and/or

brain genomic DNA was isolated and partially digested

with a combination of EcorI and EcorI Methylase. Size

selected DNA was cloned into the pBACE3.6 vector at the

ECORI sites. The ligation products were transformed into

DH10B electrocompetent cells (BRL Life Technologies)."

BASE COUNT 53 a 52 c 33 g 61 t

ORIGIN

Query Match 100.0%; Score 11; DB 13; Length 199;

Best Local Similarity 100.0%; Pred. No. 7.4e+03;

Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GCCTTAAGGC 11

|||||

Db 36 GCCTTAAGGC 26

Search completed: January 17, 2002, 08:54:03
Job time: 7115 sec

BASE COUNT 2917 a 2760 c 2957 g 3161 t
ORIGIN

Query Match 31.9%; Score 454.8; DB 6; Length 11795;
Best Local Similarity 99.6%; Pred. No. 3.4e-117;
Matches 456; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 712 GTGTAGCGTGTAGGCGCACCTTCAAGAACTCTGTAGCACCGCTTACATACCTCGCT 771
DB 986 GTGTAGCGGTAGTGTAGGCGCACCTTCAAGAACTCTGTAGCACCGCTTACATACCTCGCT 927
QY 772 CTGCTAATCTGTGTACCAAGTGGCTGTGCCAGTGGCGGATAAGTCTGTCTTACCGGGTTG 831
DB 926 CTGCTAATCTGTGTACCAAGTGGCTGTGCCAGTGGCGGATAAGTCTGTCTTACCGGGTTG 867
QY 832 GACTCAAGACCATAGTTACCGGATAAGGCGCAGCGGTGGCTGAACGGGGGGTTCGTGC 891
DB 866 GACTCAAGACCATAGTTACCGGATAAGGCGCAGCGGTGGCTGAACGGGGGGTTCGTGC 807
QY 892 ACACAGCCAGCTTGGAGCGAAGCACTTACACCGAACTGAGATACCTACAGCGTGAGCAT 951
DB 806 ACACAGCCAGCTTGGAGCGAAGCACTTACACCGAACTGAGATACCTACAGCGTGAGCAT 747
QY 952 TGAGAAAGCGCCAGCTTCCCGAAGGAGAAAGCGGACAGGTATCCGGTAAAGCGGCAGG 1011
DB 746 TGAGAAAGCGCCAGCTTCCCGAAGGAGAAAGCGGACAGGTATCCGGTAAAGCGGCAGG 687
QY 1012 GTCGGAACAGAGAGCGCAGGAGGAGCTTCCAGGGGAAACGGCTGTATCTTTATAGT 1071
DB 586 GTCGGAACAGAGAGCGCAGGAGGAGCTTCCAGGGGAAACGGCTGTATCTTTATAGT 627
QY 1072 CCTGTCCGGTTTCCGCCACCTCTGACTTGAGCGTGCATTTTGTGATGCTCGTCAGGGGG 1131
DB 626 CCTGTCCGGTTTCCGCCACCTCTGACTTGAGCGTGCATTTTGTGATGCTCGTCAGGGGG 567
QY 1132 CGGAGCCTATGGAAGAAAGCGCAGCAAGCGCGCGGGGG 1169
DB 566 CGGAGCCTATGGAAGAAAGCGCAGCAAGCGCGCGGGGG 529

RESULT 2
AR038307/c
LOCUS AR038307 13254 bp DNA PAT 29-SEP-1999
DEFINITION Sequence 156 from patent US 5804440.
ACCESSION AR038307
VERSION AR038307.1 GI:5957024
KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.
REFERENCE 1 (bases 1 to 13254)
AUTHORS Burton,D.R., Barbas,C.F. and Lerner,R.A.
TITLE Human neutralizing monoclonal antibodies to human immunodeficiency virus
JOURNAL Patent: US 5804440-A 156 08-SEP-1998;
FEATURES Location/Qualifiers
source l. 13254
BASE COUNT 3206 a 3559 c 3251 g 3238 t
ORIGIN

Query Match 31.9%; Score 454.8; DB 6; Length 13254;
Best Local Similarity 99.6%; Pred. No. 3.4e-117;
Matches 456; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 712 GTGTAGCGGTAGTGTAGGCGCACCTTCAAGAACTCTGTAGCACCGCTTACATACCTCGCT 771
DB 6449 GTGTAGCGGTAGTGTAGGCGCACCTTCAAGAACTCTGTAGCACCGCTTACATACCTCGCT 6390
QY 772 CTGCTAATCTGTGTACCAAGTGGCTGTGCCAGTGGCGGATAAGTCTGTCTTACCGGGTTG 831

DB 6389 CTGCTAATCTGTGTACCAAGTGGCTGTGCCAGTGGCGGATAAGTCTGTCTTACCGGGTTG 6330
QY 832 GACTCAAGACCATAGTTACCGGATAAGGCGCAGCGGTTCGGGTGAACGGGGGGTTCGTGC 891
DB 6329 GACTCAAGACCATAGTTACCGGATAAGGCGCAGCGGTTCGGGTGAACGGGGGGTTCGTGC 6270
QY 892 ACACAGCCAGCTTGGAGCGAAGCACTTACACCGAACTGAGATACCTACAGCGTGAGCAT 951
DB 6269 ACACAGCCAGCTTGGAGCGAAGCACTTACACCGAACTGAGATACCTACAGCGTGAGCAT 6210
QY 952 TGAGAAAGCGCCAGCTTCCCGAAGGAGAAAGCGGACAGGTATCCGGTAAAGCGGCAGG 1011
DB 6209 TGAGAAAGCGCCAGCTTCCCGAAGGAGAAAGCGGACAGGTATCCGGTAAAGCGGCAGG 6150
QY 1012 GTCGGAACAGAGAGCGCAGGAGGAGCTTCCAGGGGAAACGGCTGTATCTTTATAGT 1071
DB 6149 GTCGGAACAGAGAGCGCAGGAGGAGCTTCCAGGGGAAACGGCTGTATCTTTATAGT 6090
QY 1072 CCTGTCCGGTTTCCGCCACCTCTGACTTGAGCGTGCATTTTGTGATGCTCGTCAGGGGG 1131
DB 6089 CCTGTCCGGTTTCCGCCACCTCTGACTTGAGCGTGCATTTTGTGATGCTCGTCAGGGGG 6030
QY 1132 CGGAGCCTATGGAAGAAAGCGCAGCAAGCGCGCGGGGG 1169
DB 6029 CGGAGCCTATGGAAGAAAGCGCAGCAAGCGCGCGGGGG 5992

RESULT 3
AR038321
LOCUS AR038321 13254 bp DNA PAT 29-SEP-1999
DEFINITION Sequence 170 from patent US 5804440.
ACCESSION AR038321
VERSION AR038321.1 GI:5957038
KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.
REFERENCE 1 (bases 1 to 13254)
AUTHORS Burton,D.R., Barbas,C.F. and Lerner,R.A.
TITLE Human neutralizing monoclonal antibodies to human immunodeficiency virus
JOURNAL Patent: US 5804440-A 170 08-SEP-1998;
FEATURES Location/Qualifiers
source l. 13254
BASE COUNT 3238 a 3251 c 3559 g 3206 t
ORIGIN

Query Match 31.9%; Score 454.8; DB 6; Length 13254;
Best Local Similarity 99.6%; Pred. No. 3.4e-117;
Matches 456; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 712 GTGTAGCGGTAGTGTAGGCGCACCTTCAAGAACTCTGTAGCACCGCTTACATACCTCGCT 771
DB 6806 GTGTAGCGGTAGTGTAGGCGCACCTTCAAGAACTCTGTAGCACCGCTTACATACCTCGCT 6865
QY 772 CTGCTAATCTGTGTACCAAGTGGCTGTGCCAGTGGCGGATAAGTCTGTCTTACCGGGTTG 831
DB 6866 CTGCTAATCTGTGTACCAAGTGGCTGTGCCAGTGGCGGATAAGTCTGTCTTACCGGGTTG 6925
QY 832 GACTCAAGACCATAGTTACCGGATAAGGCGCAGCGCTCGGGCTGAACGGGGGGTTCGTGC 891
DB 6926 GACTCAAGACCATAGTTACCGGATAAGGCGCAGCGCTCGGGCTGAACGGGGGGTTCGTGC 6985
QY 892 ACACAGCCAGCTTGGAGCGAAGCACTTACACCGAACTGAGATACCTACAGCGTGAGCAT 951
DB 6986 ACACAGCCAGCTTGGAGCGAAGCACTTACACCGAACTGAGATACCTACAGCGTGAGCAT 7045
QY 952 TGAGAAAGCGCCAGCTTCCCGAAGGAGAAAGCGGACAGGTATCCGGTAAAGCGGCAGG 1011
DB 7046 TGAGAAAGCGCCAGCTTCCCGAAGGAGAAAGCGGACAGGTATCCGGTAAAGCGGCAGG 7105

QY 1012 GTCGGAACAGGAGCGACAGGAGCTTCCAGGGGAAACGCTGTATCTTTATAGT 1071
|||||
Db 7106 GTCGGAACAGGAGCGACAGGAGCTTCCAGGGGAAACGCTGTATCTTTATAGT 7165
QY 1072 CCGTGTGGGTTTCCGCACTCTGACTTGAGCGTCGATTTTGTGATGCTCGTCAGGGGG 1131
|||||
Db 7166 CCGTGTGGGTTTCCGCACTCTGACTTGAGCGTCGATTTTGTGATGCTCGTCAGGGGG 7225
QY 1132 CGGAGCCTATGGAAACGCGACGCGCGCGGGG 1169
|||||
Db 7226 CGGAGCCTATGGAAACGCGACGCGCGGGG 7263
RESULT 4
LOCUS I58596/c 158596 13254 bp DNA PAT 07-OCT-1997
DEFINITION Sequence 156 from patent US 5652138.
ACCESSION I58596
VERSION I58596.1 GI:2477834
KEYWORDS
SOURCE Unknown.
ORGANISM Unclassified.
REFERENCE 1 (bases 1 to 13254)
AUTHORS Burton,D.R., Barbas,C.F. and Lerner,R.A.
TITLE Human neutralizing monoclonal antibodies to human immunodeficiency virus
JOURNAL Patent: US 5652138-A 156 29-JUL-1997;
FEATURES Location/Qualifiers
source 1..13254
BASE COUNT 3206 a 3559 c 3251 g 3238 t
ORIGIN
Query Match 31.9%; Score 454.8; DB 6; Length 13254;
Best Local Similarity 99.6%; Pred. No. 3.4e-117;
Matches 456; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
QY 712 GTGTAGCCGTAGTTAGGCGCACCTTCAAGAATCTCTAGCAGCGCTACATACCTCGCT 771
Db 6449 GTGTAGCCGTAGTTAGGCGCACCTTCAAGAATCTCTAGCAGCGCTACATACCTCGCT 6390
QY 772 CTGCTAATCTGTACCAGTGGCTGCCAGTGGCGGCTGAAGCGGGGTTG 831
Db 6389 CTGCTAATCTGTACCAGTGGCTGCCAGTGGCGGCTGAAGCGGGGTTG 6330
QY 832 GACTCAAGACGATAGTTACCGGATAAGCGCGAGCGGCTGAAGCGGGGTTG 891
Db 6329 GACTCAAGACGATAGTTACCGGATAAGCGCGAGCGGCTGAAGCGGGGTTG 6270
QY 892 ACAGCGCCAGCTTGGAGCGAAGCACTACACCGAATGAGATACCTACAGCGTGAGCAT 951
Db 6269 ACAGCGCCAGCTTGGAGCGAAGCACTACACCGAATGAGATACCTACAGCGTGAGCAT 6210
QY 952 TGAGAAAGCGCCAGCTTCCGAGGAGAGAAAGCGGCGGCTGAAGCGGCGAG 1011
Db 6209 TGAGAAAGCGCCAGCTTCCGAGGAGAGAAAGCGGCGGCTGAAGCGGCGAG 6150
QY 1012 GTCGGAACAGGAGCGACGAGGAGCTTCCAGGGGAAACGCTGTATCTTTATAGT 1071
Db 6149 GTCGGAACAGGAGCGACGAGGAGCTTCCAGGGGAAACGCTGTATCTTTATAGT 6090
QY 1072 CCGTGTGGGTTTCCGCACTCTGACTTGAGCGTCGATTTTGTGATGCTCGTCAGGGGG 1131
Db 6089 CCGTGTGGGTTTCCGCACTCTGACTTGAGCGTCGATTTTGTGATGCTCGTCAGGGGG 6030
QY 1132 CGGAGCCTATGGAAACGCGACGCGCGGGG 1169
Db 6029 CGGAGCCTATGGAAACGCGACGCGCGGGG 5992

RESULT 5
LOCUS I58610 158610 13254 bp DNA PAT 07-OCT-1997
DEFINITION Sequence 170 from patent US 5652138.
ACCESSION I58610
VERSION I58610.1 GI:2477848
KEYWORDS
SOURCE Unknown.
ORGANISM Unclassified.
REFERENCE 1 (bases 1 to 13254)
AUTHORS Burton,D.R., Barbas,C.F. and Lerner,R.A.
TITLE Human neutralizing monoclonal antibodies to human immunodeficiency virus
JOURNAL Patent: US 5652138-A 170 29-JUL-1997;
FEATURES Location/Qualifiers
source 1..13254
BASE COUNT 3238 a 3251 c 3559 g 3206 t
ORIGIN
Query Match 31.9%; Score 454.8; DB 6; Length 13254;
Best Local Similarity 99.6%; Pred. No. 3.4e-117;
Matches 456; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
QY 712 GTGTAGCCGTAGTTAGGCGCACCTTCAAGAATCTCTAGCAGCGCTACATACCTCGCT 771
Db 6806 GTGTAGCCGTAGTTAGGCGCACCTTCAAGAATCTCTAGCAGCGCTACATACCTCGCT 6865
QY 772 CTGCTAATCTGTACCAGTGGCTGCCAGTGGCGGCTGAAGCGGGGTTG 831
Db 6866 CTGCTAATCTGTACCAGTGGCTGCCAGTGGCGGCTGAAGCGGGGTTG 6925
QY 832 GACTCAAGACGATAGTTACCGGATAAGCGCGAGCGGCTGAAGCGGGGTTGTCGTGC 891
Db 6926 GACTCAAGACGATAGTTACCGGATAAGCGCGAGCGGCTGAAGCGGGGTTGTCGTGC 6985
QY 892 ACAGCGCCAGCTTGGAGCGAAGCACTACACCGAATGAGATACCTACAGCGTGAGCAT 951
Db 6986 ACAGCGCCAGCTTGGAGCGAAGCACTACACCGAATGAGATACCTACAGCGTGAGCAT 7045
QY 952 TGAGAAAGCGCCAGCTTCCGAGGAGAGAAAGCGGCGGCTGAAGCGGCGAGG 1011
Db 7046 TGAGAAAGCGCCAGCTTCCGAGGAGAGAAAGCGGCGGCTGAAGCGGCGAGG 7105
QY 1012 GTCGGAACAGGAGCGACGAGGAGCTTCCAGGGGAAACGCTGTATCTTTATAGT 1071
Db 7106 GTCGGAACAGGAGCGACGAGGAGCTTCCAGGGGAAACGCTGTATCTTTATAGT 7165
QY 1072 CCGTGTGGGTTTCCGCACTCTGACTTGAGCGTCGATTTTGTGATGCTCGTCAGGGGG 1131
Db 7166 CCGTGTGGGTTTCCGCACTCTGACTTGAGCGTCGATTTTGTGATGCTCGTCAGGGGG 7225
QY 1132 CGGAGCCTATGGAAACGCGACGCGCGGGG 1169
Db 7226 CGGAGCCTATGGAAACGCGACGCGCGGGG 7263
RESULT 6
LOCUS AX180726/c 180726 6649 bp DNA PAT 06-AUG-2001
DEFINITION Sequence 6 from Patent WO0146694.
ACCESSION AX180726
VERSION AX180726.1 GI:15132581
KEYWORDS
SOURCE synthetic construct.
ORGANISM synthetic construct.
REFERENCE 1 (bases 1 to 6649)
AUTHORS Joly,E.
TITLE A bioluminescence resonance energy transfer (bret) fusion molecule and method of use

Fri Jan 18 08:27:36 2002

JOURNAL Patent: WO 014694-A 6 28-JUN-2001;
Biosignal
Location/Qualifiers
1. .6649
Source /organism="synthetic construct"
/db_xref="taxon:32630"
/note="DNA sequence for Rluc-PKA-EYFP construct"

BASE COUNT 1632 a 1685 g 1647 t
ORIGIN

Query Match 31.9%; Score 454.2; DB 6; Length 6649;
Best Local Similarity 98.3%; Pred. No. 5e-117;
Matches 459; Conservative 0; Mismatches 8; Indels 0; Gaps 0;

QY 698 AATTCCTGAGCCGGTGTAGCCCTAGTTAGGCCACACCTTCRAGAACTCTGTAGCACCGC 757
DB 5343 AATCTGCTCTTAGTGTAGCCGTAGTTAGGCCACACCTTCRAGAACTCTGTAGCACCGC 5284
QY 758 CTACATACCTCGCTCTCTCTAATCTGTTTACCAGTGGCTGCTGCCAGTGGCGATAAGTCGT 817
DB 5283 CTACATACCTCGCTCTCTCTAATCTGTTTACCAGTGGCTGCTGCCAGTGGCGATAAGTCGT 5224
QY 818 GTCTTACCGGGTTGGACTCAAGACGATAGTTACCGGATAAGGCGCAGCGTCCGGCTGAA 877
DB 5223 GTCTTACCGGGTTGGACTCAAGACGATAGTTACCGGATAAGGCGCAGCGTCCGGCTGAA 5164
QY 878 CGGGGGTTCGTGCACACAGCCAGCTTGGAGCGAAGACCTACACCGAACTGAGATACC 937
DB 5163 CGGGGGTTCGTGCACACAGCCAGCTTGGAGCGAAGACCTACACCGAACTGAGATACC 5104
QY 938 TACAGCGTGAGCAITGAGAAAGCCACGCTTCCCGAAGGAGAAAGCGGACAGGTATC 997
DB 5103 TACAGCGTGAGCAITGAGAAAGCCACGCTTCCCGAAGGAGAAAGCGGACAGGTATC 5044
QY 998 CGGTAAAGCGGAGGTTCGACACAGGAGAGCGACAGGAGCTTCCAGGGGAAACCGCT 1057
DB 5043 CGGTAAAGCGGAGGTTCGACACAGGAGAGCGACAGGAGCTTCCAGGGGAAACCGCT 4984
QY 1058 GGTATCTTTATAGTCTCTCGGGTTTCGCCACCTCTGACTTGAGCGTGGATTTTGTGAT 1117
DB 4983 GGTATCTTTATAGTCTCTCGGGTTTCGCCACCTCTGACTTGAGCGTGGATTTTGTGAT 4924
QY 1118 GCTGCTAGGGGGGGAGCCCTATGGAAGAAAGCGCAGCAAGCGGCC 1164
DB 4923 GCTGCTAGGGGGGGAGCCCTATGGAAGAAAGCGCAGCAAGCGGCC 4877

RESULT 7
AR027070/c AR027070 1905 bp DNA PAT 29-SEP-1999
LOCUS Sequence 9 from patent US 5856142.
DEFINITION AR027070
ACCESSION AR027070
VERSION AR027070.1 GI:5937910
KEYWORDS
SOURCE Unknown.
ORGANISM Unclassified.
REFERENCE 1 (bases 1 to 1905)
AUTHORS Legoux,R., Maldonado,P. and Salome,M.
TITLE Method for the extraction of periplasmic proteins from prokaryotic microorganisms in the presence of arginine
JOURNAL Patent: US 5856142-A 9 05-JAN-1999;
FEATURES Location/Qualifiers
source 1. .1905
BASE COUNT 479 a 493 c 442 g 491 t
ORIGIN

Query Match 31.8%; Score 453.8; DB 6; Length 1905;
Best Local Similarity 99.6%; Pred. No. 6.5e-117;
Matches 455; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 712 GTGTAGCCGTAGTTAGGCCACCACTTCAAGAACTCTGTAGCACCGCTACATACCTCGCT 771
DB 518 GTGTAGCCGTAGTTAGGCCACCACTTCAAGAACTCTGTAGCACCGCTACATACCTCGCT 459
QY 772 CTGCTAATCTCTGTACCAGTGGCTGCTGCCAGTGGCGATAAGTCTGTCTTACCGGGTTG 831
DB 458 CTGCTAATCTCTGTACCAGTGGCTGCTGCCAGTGGCGATAAGTCTGTCTTACCGGGTTG 399
QY 832 GACTCAAGACGATAGTTACCGGATAAGCGGAGCGGTCGGGCTGAACGGGGGTTCTGTGC 891
DB 398 GACTCAAGACGATAGTTACCGGATAAGCGGAGCGGTCGGGCTGAACGGGGGTTCTGTGC 339
QY 892 ACACAGCCCACTTGGAGCGAACCTACACCGAACTGAGATACCTACACGCGTGAGCAT 951
DB 338 ACACAGCCCACTTGGAGCGAACCTACACCGAACTGAGATACCTACACGCGTGAGCAT 279
QY 952 TGAGAAAGCGCCAGCTTCCCGAAGGGAGAAAGCGGACAGGTATCCGGTAAGCGGCAGG 1011
DB 278 TGAGAAAGCGCCAGCTTCCCGAAGGGAGAAAGCGGACAGGTATCCGGTAAGCGGCAGG 219
QY 1012 GTGCGAAGCAGAGAGCGCACGAGGAGCTTCCAGGGGAAACGCTGTATCTTTATAGT 1071
DB 218 GTGCGAAGCAGAGAGCGCACGAGGAGCTTCCAGGGGAAACGCTGTATCTTTATAGT 159
QY 1072 CTTGTCTGGGTTTTCGCCACCTCTGACTTGAGCGTTCGATTTTGTGATGCTCTCAGGGGG 1131
DB 158 CTTGTCTGGGTTTTCGCCACCTCTGACTTGAGCGTTCGATTTTGTGATGCTCTCAGGGGG 99
QY 1132 CGGAGCCTATGGAAGAAAGCGCAGCAAGCGGCCGGG 1168
DB 98 CGGAGCCTATGGAAGAAAGCGCAGCAAGCGGCCGGG 62

RESULT 8
I86203/c I86203 1905 bp DNA PAT 10-JUN-1998
LOCUS Sequence 9 from patent US 5700665.
DEFINITION I86203
ACCESSION I86203
VERSION I86203.1 GI:3205921
KEYWORDS
SOURCE Unknown.
ORGANISM Unclassified.
REFERENCE 1 (bases 1 to 1905)
AUTHORS Legoux,R., Maldonado,P. and Salome,M.
TITLE Method for the extraction of periplasmic proteins from prokaryotic microorganisms in the presence of arginine
JOURNAL Patent: US 5700665-A 9 23-DEC-1997;
FEATURES Location/Qualifiers
source 1. .1905
BASE COUNT 479 a 493 c 442 g 491 t
ORIGIN

Query Match 31.8%; Score 453.8; DB 6; Length 1905;
Best Local Similarity 99.6%; Pred. No. 6.5e-117;
Matches 455; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 712 GTGTAGCCGTAGTTAGGCCACCACTTCAAGAACTCTGTAGCACCGCTACATACCTCGCT 771
DB 518 GTGTAGCCGTAGTTAGGCCACCACTTCAAGAACTCTGTAGCACCGCTACATACCTCGCT 459
QY 772 CTGCTAATCTCTGTACCAGTGGCTGCTGCCAGTGGCGATAAGTCTGTCTTACCGGGTTG 831
DB 458 CTGCTAATCTCTGTACCAGTGGCTGCTGCCAGTGGCGATAAGTCTGTCTTACCGGGTTG 399
QY 832 GACTCAAGACGATAGTTACCGGATAAGCGGAGCGGTCGGGCTGAACGGGGGTTCTGTGC 891
DB 398 GACTCAAGACGATAGTTACCGGATAAGCGGAGCGGTCGGGCTGAACGGGGGTTCTGTGC 339
QY 892 ACACAGCCCACTTGGAGCGAACGACCTACACCGAACTGAGATACCTACACGCGTGAGCAT 951


```
|||||
Db 338 ACACAGCCAGCTTGGAGGAAACACCTTACACCGAACTGAGATACCTACAGCGTGAGCAT 279
Qy 952 TGAGAAAGCGCCACGCTTCCCGAGGAGGAGAAAGCGGACAGGTATCCGGTAAGCGGCAGG 1011
Db 278 TGAGAAAGCGCCACGCTTCCCGAGGAGGAGGAGAAAGCGGACAGGTATCCGGTAAGCGGCAGG 219
Qy 1012 GTCGGAACAGGAGAGCGCACGAGGAGGAGCTTCCAGGGGAAACGCGCTGATCTTTATAGT 1071
Db 218 GTCGGAACAGGAGAGCGCACGAGGAGGAGCTTCCAGGGGAAACGCGCTGATCTTTATAGT 159
Qy 1072 CCTGTCGGGTTTCGCCACCTCTGACTTGAGCGTGCATTTTGTGATGCTCTGTCAGGGGGG 1131
Db 158 CTTGTCGGGTTTCGCCACCTCTGACTTGAGCGTGCATTTTGTGATGCTCTGTCAGGGGGG 99
Qy 1132 CGGAGCCTATGGAAGAAAGCGGAGCAACGCGCGCGGG 1168
Db 98 CGGAGCCTATGGAAGAAAGCGGAGCAACGCGCGCGGAG 62

RESULT 9
ARP327/c ARP327 3273 bp DNA circular SYN 17-MAY-1995
LOCUS Artificial cloning vehicle pBR327, derived from pBR322. The
DEFINITION sequence was not resequenced but deduced from the pBR322 sequence.
Contains the reading frames for ampicillin resistance (Apr) and
tetracycline resistance (Tcr) and an origin of replication.
ACCESSION V00083 J02549
VERSION V00083.1 GI:58119
KEYWORDS circular; cloning vector; drug resistance gene; origin of
SOURCE replication; plasmid.
ORGANISM synthetic construct.
ARTIFICIAL SEQUENCE.
REFERENCE 1 (bases 1 to 3273)
AUTHORS Soberon,X., Covarrubias,L. and Bolivar,F.
TITLE Construction and characterization of new cloning vehicles. IV.
Deletion derivatives of pBR322 and pBR325
JOURNAL Gene 9 (3-4), 287-305 (1980)
MEDLINE 80225760
FEATURES Location/Qualifiers
source 1..3273
/organism="synthetic construct"
/db_xref="taxon:32630"
259..1275
/note="tetracycline resistance"
/codon_start=1
/transl_table=11
/protein_id="CAA23425.1"
/db_xref="GI:808934"
/transl_table="MRTRSRSTVRPLWPPSPARFATWSHVRLRDHGHDPVDPPLRR
THGRHRRHRCRCWRLYRRHRWGRSGPLRAHERLFRGCGRPVAGGLGALSLH
APFLAAVNLNLLGCFLMQSHKGERPMLRAENPVSSFWARGMTIVAAIATV
FFIMLVQVQPAALVTFGEDRFWSATMIGLSLAVFGLHALAQAEVTPATKRFGE
KQAIAGMAADALGYLLAFATPAGWMAFPIIMILASGICMPALQMLSRQVDDHOG
QLQGLAALSLTSITGPLIVTATYAAASASTWNGIAMIWGAALVYLCLPALRRGAWSR
ATST"
complement(2205..3065)
/note="ampicillin resistance"
/codon_start=1
/transl_table=11
/protein_id="CAA23426.1"
/db_xref="GI:808935"
/transl_table="MSIQHFVVALIPFAAFCLPVFAHPETLVKVKDAEDQLGARVGY
IEDLNKGLDESFRPEERPMSTFKVLICGAVLSRVDAQQEOLGRIIRHYSQNDLVE
YSPTEKHTLDIVRELCASATNTAANLLLTIGGPKELTAPLHNMGDHTRL
DRWEPELNEAIPINDERDTPMAAMATLRKLTLGELLTLASRQOLIDWMEADKVGPL
LRGALPAGWTFIADKSGAGERSGILIAALPGDGPSPRIVVIYITGSGATWDERNRQIA
EIGASLLKHW"
BASE COUNT 721 a 911 c 835 g 806 t
ORIGIN
```

```
Query Match 31.8%; Score 453.8; DB 12; Length 3273;
Best Local Similarity 99.6%; Pred. No. 6.5e-117;
Matches 455; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 712 GTGTAGCGGTAGTTAGGCCACCACTTCAAGAAGTCTGTAGCACCGCCCTACATACCTCGCT 771
Db 1880 GTGTAGCGGTAGTTAGGCCACCACTTCAAGAAGTCTGTAGCACCGCCCTACATACCTCGCT 1821
Qy 772 CTGCTAATCTGTTTACCAGTGCCTGCCAGTGGCGGATAAGTCTGCTTTACCGGGTTG 831
Db 1820 CTGCTAATCTGTTTACCAGTGCCTGCCAGTGGCGGATAAGTCTGCTTTACCGGGTTG 1761
Qy 832 GACTCAAGACGATAGTTACCAGGATAAGCGCGAGCGGTCCGGCTGAACGGGGGGTTTCGTGC 891
Db 1760 GACTCAAGACGATAGTTACCAGGATAAGCGCGAGCGGTCCGGCTGAACGGGGGGTTTCGTGC 1701
Qy 892 ACACAGCCCGAGCTTGGAGCGAAGACCTACACCGAAGTACCTACAGCGTGAGCAT 951
Db 1700 ACACAGCCCGAGCTTGGAGCGAAGACCTACACCGAAGTACCTACAGCGTGAGCAT 1641
Qy 952 TGAGAAAGCGCCACGCTTCCCGAGGAGGAGAAAGCGGACAGGTATCCGTAAGCGGCAGG 1011
Db 1640 TGAGAAAGCGCCACGCTTCCCGAGGAGGAGAAAGCGGACAGGTATCCGTAAGCGGCAGG 1581
Qy 1012 GTCGGAACAGGAGAGCGCACGAGGAGGAGCTTCCAGGGGAAACGCGCTGATCTTTATAGT 1071
Db 1580 GTCGGAACAGGAGAGCGCACGAGGAGGAGCTTCCAGGGGAAACGCGCTGATCTTTATAGT 1521
Qy 1072 CCTGTCGGGTTTCGCCACCTCTGACTTGAGCGTGCATTTTGTGATGCTCTGTCAGGGGGG 1131
Db 1520 CTTGTCGGGTTTCGCCACCTCTGACTTGAGCGTGCATTTTGTGATGCTCTGTCAGGGGGG 1461
Qy 1132 CGGAGCCTATGGAAGAAAGCGGAGCAACGCGCGCGGG 1168
Db 1460 CGGAGCCTATGGAAGAAAGCGGAGCAACGCGCGCGGAG 1424

RESULT 10
A49702/c A49702 4023 bp DNA circular PAT 07-MAR-1997
LOCUS Sequence 8 from Patent WO9610641.
DEFINITION A49702
ACCESSION A49702
VERSION A49702.1 GI:2303038
KEYWORDS Influenza virus.
SOURCE Influenza virus.
ORGANISM Viruses; ssRNA negative-strand viruses; Orthomyxoviridae;
unclassified Orthomyxoviridae.
REFERENCE 1 (bases 1 to 4023)
AUTHORS Hobom,G., Neumann,G. and Menke,A.
TITLE An attenuated vaccination virus, a method to make the virus and a
pharmaceutical compositions comprising the virus
JOURNAL Patent: WO 9610641-A 8 11-APR-1996;
COMMENT BAYER AG (DE)
FEATURES Other publication AU 3607695 960426.
Location/Qualifiers
source 1..4023
/organism="Influenza virus"
/isolate="PHLI490"
/db_xref="taxon:11309"
BASE COUNT 1071 a 1001 c 1004 g 947 t
ORIGIN
```

```
Query Match 31.8%; Score 453.8; DB 6; Length 4023;
Best Local Similarity 99.6%; Pred. No. 6.5e-117;
Matches 455; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 712 GTGTAGCGGTAGTTAGGCCACCACTTCAAGAAGTCTGTAGCACCGCCCTACATACCTCGCT 771
Db 2094 GTGTAGCGGTAGTTAGGCCACCACTTCAAGAAGTCTGTAGCACCGCCCTACATACCTCGCT 2035
Qy 772 CTGCTAATCTGTTTACCAGTGCCTGCCAGTGGCGGATAAGTCTGCTTTACCGGGTTG 831
```

Db	2034	CTGCTAAATCTCTTACCACTGCTCTCCAGTGGCGATAAGTCGGTGTCTTACCGGGTTG	1975
QY	832	GACTCAAGACGATAGTTACCGGATAAGGCGCAGCGTGGGCTGAACGGGGGTTTCGTGC	891
Db	1974	GACTCAAGACGATAGTTACCGGATAAGGCGCAGCGTGGGCTGAACGGGGGTTTCGTGC	1915
QY	892	ACACAGCCCGCTTGGAGGACGACCTACACCGAAGTACGATACCTACAGCGTGAGCAT	951
Db	1914	ACACAGCCCGCTTGGAGGACGACCTACACCGAAGTACGATACCTACAGCGTGAGCAT	1855
QY	952	TGAGAAAGCGCCACGCTTCCCGAAGGGAGAAAGCGGACAGGTATCCGGTAAGCGCGAGG	1011
Db	1854	TGAGAAAGCGCCACGCTTCCCGAAGGGAGAAAGCGGACAGGTATCCGGTAAGCGCGAGG	1795
QY	1012	GTCGAAACAGGAGACGCGACGAGGAGGCTTCAGGGGGAAACGCTTGTATCTTTATAGT	1071
Db	1794	GTCGAAACAGGAGACGCGACGAGGAGGCTTCAGGGGGAAACGCTTGTATCTTTATAGT	1735
QY	1072	CCTGTCGGGTTTGGCCACCTCTGACTTCAGCGTTCATTTTGTGATGCTCGTCAGGGGG	1131
Db	1734	CCTGTCGGGTTTGGCCACCTCTGACTTCAGCGTTCATTTTGTGATGCTCGTCAGGGGG	1675
QY	1132	CGGAGCCTATGAAAAACCCAGCAACGCGCGCGGG	1168
Db	1674	CGGAGCCTATGAAAAACCCAGCAACGCGCGCGGAG	1638
RESULT	11		
XXU13846/c			
LOCUS	XXU13846	4060 bp	DNA circular SYN 13-DEC-1994
DEFINITION	pCM7 cloning vector, complete sequence.		
ACCESSION	U13846		
VERSION	U13846.1 GI:595696		
KEYWORDS	chloramphenicol acetyltransferase; promoter analysis; beta-lactamase.		
SOURCE	unidentified cloning vector.		
ORGANISM	unidentified cloning vector		
REFERENCE	artificial sequence; vectors.		
AUTHORS	1 (bases 1 to 4060)		
TITLE	Malone,J.A.		
JOURNAL	pCM7: A promoter analysis, chloramphenicol acetyltransferase gene		
REFERENCE	cartridge vector		
AUTHORS	Unpublished (1994)		
TITLE	2 (bases 1 to 4060)		
JOURNAL	Close,T.J. and Rodriguez,R.L.		
REFERENCE	Construction and characterization of the chloramphenicol-resistance		
AUTHORS	gene cartridge: a new approach to the transcriptional mapping of		
TITLE	extrachromosomal elements		
JOURNAL	Gene 20 (2), 305-316 (1982)		
MEDLINE	83158767		
REFERENCE	3 (bases 38 to 812)		
AUTHORS	Alton,N.K. and Vapnek,D.		
TITLE	Nucleotide sequence analysis of the chloramphenicol resistance		
JOURNAL	transposon Tn9		
MEDLINE	Nature 282 (5741), 864-869 (1979)		
REFERENCE	80078150		
AUTHORS	4 (bases 1 to 4060)		
TITLE	Malone,J.A.		
JOURNAL	Direct Submission		
REFERENCE	Submitted (19-AUG-1994) James A. Malone, International Technical		
AUTHORS	Services, Molecular Biology Reagents Division, Pharmacia Biotech		
TITLE	Inc., 2202 N. Bartlett Ave., Milwaukee, WI 53202-1009, USA		
FEATURES	Location/Qualifiers		
source	1..4060		
	/organism="unidentified cloning vector"		
	/db_xref="taxon:45196"		
gene	/lab_host="Escherichia coli"		
	67..726		
	/gene="cat"		
CDS	67..726		
	/gene="cat"		
	/note="The cat gene in pCM7 is promoterless The cat gene		

RESULT 12
SYNPBR329/c
LOCUS SYNPBR329 4150 bp DNA circular SYN 27-APR-1993
DEFINITION Plasmid pBR329, complete genome.
ACCESSION J01753
VERSION J01753.1 GI:208964
KEYWORDS cloning vector; complete genome.
SOURCE plasmid pbr329 from E.coli k-12.
ORGANISM unidentified cloning vector
artificial sequence; vectors.
REFERENCE 1 (bases 1 to 4150)
AUTHORS Covarrubias,L. and Bolivar,F.
TITLE Construction and characterization of new cloning vehicles: Vi.
Plasmid pBR329, a new derivative of pBR328 lacking the
482-base-pair inverted duplication
JOURNAL Gene 17, 79-89 (1982)
MEDLINE 8218782
COMMENT [1] only sequenced bases 160-340 and 3440-3560. the other regions
are from previously published sequences for pBR327 and Tn981 (see
fig. 5 for refs).
FEATURES
source Location/Qualifiers
1. .4150
/organism="unidentified cloning vector"
/db_xref="taxon:45196"
BASE COUNT 970 a 1121 c 1021 g 1038 t
ORIGIN
Query Match 31.8%; Score 453.8; DB 12; Length 4150;
Best Local Similarity 99.6%; Pred. No. 6.5e-117;
Matches 455; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
Qy 712 GTGTAGCGGTAGTTAGGCCACCACCTTCAAGAACTCTGTAGCACCGCCTACATACCTCGCT 771
Db 2111 GTGTAGCGGTAGTTAGGCCACCACCTTCAAGAACTCTGTAGCACCGCCTACATACCTCGCT 2052
Qy 772 CTGCTAATCCTGTTACCAAGTGGCTGCCAGTGGCGGATAAGTGTGTTTACCGGGTTG 831
Db 2051 CTGCTAATCCTGTTACCAAGTGGCTGCCAGTGGCGGATAAGTGTGTTTACCGGGTTG 1992
Qy 832 GACTCAAGACGATAGTTACCGGATAAGCGCAGCGGTGGCTGAAACGGGGGTTTCGTGC 891
Db 1991 GACTCAAGACGATAGTTACCGGATAAGCGCAGCGGTGGCTGAAACGGGGGTTTCGTGC 1932
Qy 892 ACACAGCCACCTTGGACGCAACACCTACACCGAAGTACGATACCTACACGCGTGAGCAT 951
Db 1931 ACACAGCCACCTTGGACGCAACACCTACACCGAAGTACGATACCTACACGCGTGAGCAT 1872
Qy 952 TGAGAAAGCGCCACGCTTCCCGAAGGGAGAAAGCGGACAGGTATCCGGTAAGCGGCAGG 1011
Db 1871 TGAGAAAGCGCCACGCTTCCCGAAGGGAGAAAGCGGACAGGTATCCGGTAAGCGGCAGG 1812
Qy 1012 GTCGGAACAGGAGCGCACGAGGAGGCTTCCAGGGGAAACGCTTGATCTTTATAGT 1071
Db 1811 GTCGGAACAGGAGCGCACGAGGAGGCTTCCAGGGGAAACGCTTGATCTTTATAGT 1752
Qy 1072 CCTGTCGGGTTTCGCCACCTCTGACTTGAGCGTCGATTTTGTGATGCTCGTCAGGGGGG 1131
Db 1751 CCTGTCGGGTTTCGCCACCTCTGACTTGAGCGTCGATTTTGTGATGCTCGTCAGGGGGG 1692
Qy 1132 CGGAGCCTATGAAAGCGCCAGCAACGCGCCGGG 1168
Db 1691 CGGAGCCTATGAAAGCGCCAGCAACGCGCCGGG 1655
RESULT 13
A52326/c
LOCUS A52326 4410 bp DNA PAT 12-DEC-1997
DEFINITION Sequence 1 from Patent EP0725140.
ACCESSION A52326
VERSION A52326.1 GI:2851987

KEYWORDS unidentified.
SOURCE unidentified.
ORGANISM unclassified.
REFERENCE 1 (bases 1 to 4410)
AUTHORS Legoux,R., Maldonado,P. and Salome,M.
TITLE Process of extraction of periplasmic proteins from prokaryotic
microorganisms in the presence of arginine
JOURNAL Patent: EP 0725140-A 1 07-AUG-1996;
SANOFI SA (FR)
COMMENT Other publication SK 10696 960904
Other publication Cz 9600290 960814
Other publication JP 8242879 960924
Other publication FI 960427 960801
Other publication PL 312543 960805
Other publication NO 960396 960801
Other publication FR 2729372 960802
Other publication CA 2168382 960801
Other publication AU 4224496 960808.
FEATURES
source Location/Qualifiers
1. .4410
/organism="unidentified"
/db_xref="taxon:32644"
1. .282
283. .337
338. .762
/note="SEQUENCE DE LA REGION 5' NON TRADUITE DU MESSAGE"
misc_RNA
763. .812
/note="SEQUENCE CODANT POUR LE PRECURSEUR DE L'IL-13"
terminator
813. .1012
/note="TERMINATEUR DU GENE 10 DU PHAGE T7"
terminator
1013. .1253
/note="TERMINATEUR DU PHAGE FD"
misc_RNA
1254. .2505
/note="GENE CODANT POUR LE REPRESSEUR DE L'OPERON LACTOSE"
misc_RNA
2506. .4410
/note="SEQUENCE DE PBR 327"
BASE COUNT 1078 a 1142 c 1096 g 1094 t
ORIGIN
Query Match 31.8%; Score 453.8; DB 6; Length 4410;
Best Local Similarity 99.6%; Pred. No. 6.5e-117;
Matches 455; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
Qy 712 GTGTAGCGGTAGTTAGGCCACCACCTTCAAGAACTCTGTAGCACCGCCTACATACCTCGCT 771
Db 3023 GTGTAGCGGTAGTTAGGCCACCACCTTCAAGAACTCTGTAGCACCGCCTACATACCTCGCT 2964
Qy 772 CTGCTAATCCTGTTACCAAGTGGCTGCCAGTGGCGGATAAGTGTGTTTACCGGGTTG 831
Db 2963 CTGCTAATCCTGTTACCAAGTGGCTGCCAGTGGCGGATAAGTGTGTTTACCGGGTTG 2904
Qy 832 GACTCAAGACGATAGTTACCGGATAAGCGCAGCGGTGGCTGAAACGGGGGTTTCGTGC 891
Db 2903 GACTCAAGACGATAGTTACCGGATAAGCGCAGCGGTGGCTGAAACGGGGGTTTCGTGC 2844
Qy 892 ACACAGCCACCTTGGACGCAACACCTACACCGAAGTACGATACCTACAGCGTGAGCAT 951
Db 2843 ACACAGCCACCTTGGACGCAACACCTACACCGAAGTACGATACCTACAGCGTGAGCAT 2784
Qy 952 TGAGAAAGCGCCACGCTTCCCGAAGGGAGAAAGCGGACAGGTATCCGGTAAGCGGCAGG 1011
Db 2783 TGAGAAAGCGCCACGCTTCCCGAAGGGAGAAAGCGGACAGGTATCCGGTAAGCGGCAGG 2724
Qy 1012 GTCGGAACAGGAGCGCACGAGGAGGCTTCCAGGGGAAACGCTTGATCTTTATAGT 1071
Db 2723 GTCGGAACAGGAGCGCACGAGGAGGCTTCCAGGGGAAACGCTTGATCTTTATAGT 2664
Qy 1072 CCTGTCGGGTTTCGCCACCTCTGACTTGAGCGTCGATTTTGTGATGCTCGTCAGGGGGG 1131
Db 2663 CCTGTCGGGTTTCGCCACCTCTGACTTGAGCGTCGATTTTGTGATGCTCGTCAGGGGGG 2604

```
QY 1132 CGGACCTATGGAAGAACCCAGCAACCCGCGCGGG 1168
|||||
Db 2603 CGGACCTATGGAAGAACCCAGCAACCCGCGCGGAG 2567

RESULT 14
AR027062/c
LOCUS AR027062 4410 bp DNA PAT 29-SEP-1999
DEFINITION Sequence 1 from patent US 5856142.
ACCESSION AR027062
VERSION AR027062.1 GI:5937902
KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.
REFERENCE 1 (bases 1 to 4410)
AUTHORS Legoux,R., Maldonado,P. and Salome,M.
TITLE Method for the extraction of periplasmic proteins from prokaryotic
microorganisms in the presence of arginine
JOURNAL Patent: US 5856142-A 1 05-JAN-1999;
FEATURES
Location/Qualifiers
Location 1..4410
BASE COUNT 1078 a 1142 c 1096 g 1094 t
ORIGIN

Query Match 31.8%; Score 453.8; DB 6; Length 4410;
Best Local Similarity 99.8%; Pred. No. 6.5e-117;
Matches 455; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 712 GTGTAGCCGTAGTTAGGCCACCACTTCAAGAACTCTGTAGCAGCGCTACATACCTCGCT 771
Db 3023 GTGTAGCCGTAGTTAGGCCACCACTTCAAGAACTCTGTAGCAGCGCTACATACCTCGCT 2964

QY 772 CTGCTAATCTGTTACCAAGTGGCTGCTGCGAGTGGCGATAAGTCTGTCTTACCGGGTTG 831
Db 2963 CTGCTAATCTGTTACCAAGTGGCTGCTGCGAGTGGCGATAAGTCTGTCTTACCGGGTTG 2904

QY 832 GACTCAAGACGATAGTTACCGGATAAGCGCGACGCTGCGGCTGAACGGGGGTTTCGTGC 891
Db 2903 GACTCAAGACGATAGTTACCGGATAAGCGCGACGCTGCGGCTGAACGGGGGTTTCGTGC 2844

QY 892 ACACAGCCAGCTTGGAGCGAGACCTACACCGAACTGAGATACCTACAGCGTGAGCAT 951
Db 2843 ACACAGCCAGCTTGGAGCGAGACCTACACCGAACTGAGATACCTACAGCGTGAGCAT 2784

QY 952 TGAGAAAGCCACGCTTCCCGAAGGAGAAAGCGGACAGGTATCCGGTAAGCGGCAGG 1011
Db 2783 TGAGAAAGCCACGCTTCCCGAAGGAGAAAGCGGACAGGTATCCGGTAAGCGGCAGG 2724

QY 1012 GTCGGAACAGGAGCGCAGCGAGGAGCTTCCAGGGGGAACGCCCTGGTATCTTTATAGT 1071
Db 2723 GTCGGAACAGGAGCGCAGCGAGGAGCTTCCAGGGGGAACGCCCTGGTATCTTTATAGT 2664

QY 1072 CCGTCCGGGTTTCGCCACCTCTGACTTGAGCGTCGATTTTGTGATGCTCGTCAGGGGG 1131
Db 2663 CCGTCCGGGTTTCGCCACCTCTGACTTGAGCGTCGATTTTGTGATGCTCGTCAGGGGG 2604

QY 1132 CGGAGCCTATGGAAGAACCCAGCAACCCGCGCGGG 1168
|||||
Db 2603 CGGAGCCTATGGAAGAACCCAGCAACCCGCGCGGAG 2567

Search completed: January 17, 2002, 11:49:01
Job time: 17153 sec
```

```
QY 1132 CGGACCTATGGAAGAACCCAGCAACCCGCGCGGG 1168
|||||
Db 2603 CGGACCTATGGAAGAACCCAGCAACCCGCGCGGAG 2567

RESULT 15
I86195/c
LOCUS I86195 4410 bp DNA PAT 10-JUN-1998
DEFINITION Sequence 1 from patent US 5700665.
ACCESSION I86195
VERSION I86195.1 GI:3205913
KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.

Query Match 31.8%; Score 453.8; DB 6; Length 4410;
Best Local Similarity 99.8%; Pred. No. 6.5e-117;
Matches 455; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 712 GTGTAGCCGTAGTTAGGCCACCACTTCAAGAACTCTGTAGCAGCGCTACATACCTCGCT 771
Db 3023 GTGTAGCCGTAGTTAGGCCACCACTTCAAGAACTCTGTAGCAGCGCTACATACCTCGCT 2964

QY 772 CTGCTAATCTGTTACCAAGTGGCTGCTGCGAGTGGCGATAAGTCTGTCTTACCGGGTTG 831
Db 2963 CTGCTAATCTGTTACCAAGTGGCTGCTGCGAGTGGCGATAAGTCTGTCTTACCGGGTTG 2904

QY 832 GACTCAAGACGATAGTTACCGGATAAGCGCGACGCTGCGGCTGAACGGGGGTTTCGTGC 891
Db 2903 GACTCAAGACGATAGTTACCGGATAAGCGCGACGCTGCGGCTGAACGGGGGTTTCGTGC 2844

QY 892 ACACAGCCAGCTTGGAGCGAGACCTACACCGAACTGAGATACCTACAGCGTGAGCAT 951
Db 2843 ACACAGCCAGCTTGGAGCGAGACCTACACCGAACTGAGATACCTACAGCGTGAGCAT 2784

QY 952 TGAGAAAGCCACGCTTCCCGAAGGAGAAAGCGGACAGGTATCCGGTAAGCGGCAGG 1011
Db 2783 TGAGAAAGCCACGCTTCCCGAAGGAGAAAGCGGACAGGTATCCGGTAAGCGGCAGG 2724

QY 1012 GTCGGAACAGGAGCGCAGCGAGGAGCTTCCAGGGGGAACGCCCTGGTATCTTTATAGT 1071
Db 2723 GTCGGAACAGGAGCGCAGCGAGGAGCTTCCAGGGGGAACGCCCTGGTATCTTTATAGT 2664

QY 1072 CCGTCCGGGTTTCGCCACCTCTGACTTGAGCGTCGATTTTGTGATGCTCGTCAGGGGG 1131
Db 2663 CCGTCCGGGTTTCGCCACCTCTGACTTGAGCGTCGATTTTGTGATGCTCGTCAGGGGG 2604

QY 1132 CGGAGCCTATGGAAGAACCCAGCAACCCGCGCGGG 1168
|||||
Db 2603 CGGAGCCTATGGAAGAACCCAGCAACCCGCGCGGAG 2567
```


THIS PAGE BLANK (USPTO)

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: January 17, 2002, 12:02:19 ; Search time 599.86 Seconds
(without alignments)
2036.623 Million cell updates/sec

Title: US-09-242-202a-16
Perfect score: 1425
Sequence: 1 TCCCATGGCGGATCTTT.....CAGCCTCTCCACAGGTACC 1425

Scoring table: IDENTITY_NUC
Gap 10.0 , Gapext 1.0

Searched: 930621 seqs, 428662619 residues
Total number of hits satisfying chosen parameters: 1861242

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : N_Geneseq_1101.*
1: /SIDS2/gcgdata/geneseq/geneseq/NA1980.DAT.*
2: /SIDS2/gcgdata/geneseq/geneseq/NA1981.DAT.*
3: /SIDS2/gcgdata/geneseq/geneseq/NA1982.DAT.*
4: /SIDS2/gcgdata/geneseq/geneseq/NA1983.DAT.*
5: /SIDS2/gcgdata/geneseq/geneseq/NA1984.DAT.*
6: /SIDS2/gcgdata/geneseq/geneseq/NA1985.DAT.*
7: /SIDS2/gcgdata/geneseq/geneseq/NA1986.DAT.*
8: /SIDS2/gcgdata/geneseq/geneseq/NA1987.DAT.*
9: /SIDS2/gcgdata/geneseq/geneseq/NA1988.DAT.*
10: /SIDS2/gcgdata/geneseq/geneseq/NA1989.DAT.*
11: /SIDS2/gcgdata/geneseq/geneseq/NA1990.DAT.*
12: /SIDS2/gcgdata/geneseq/geneseq/NA1991.DAT.*
13: /SIDS2/gcgdata/geneseq/geneseq/NA1992.DAT.*
14: /SIDS2/gcgdata/geneseq/geneseq/NA1993.DAT.*
15: /SIDS2/gcgdata/geneseq/geneseq/NA1994.DAT.*
16: /SIDS2/gcgdata/geneseq/geneseq/NA1995.DAT.*
17: /SIDS2/gcgdata/geneseq/geneseq/NA1996.DAT.*
18: /SIDS2/gcgdata/geneseq/geneseq/NA1997.DAT.*
19: /SIDS2/gcgdata/geneseq/geneseq/NA1998.DAT.*
20: /SIDS2/gcgdata/geneseq/geneseq/NA1999.DAT.*
21: /SIDS2/gcgdata/geneseq/geneseq/NA2000.DAT.*
22: /SIDS2/gcgdata/geneseq/geneseq/NA2001.DAT.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	ID	Description
1	1425	100.0	1425	19	AAV21724 Humanised vector p
2	1203	84.4	2125	19	AAV21727 Humanised vector p
3	1194.4	83.8	1911	19	AAV21726 Humanised vector p
4	490.4	34.4	1547	19	AAV21732 Humanised vector p
5	459.8	32.3	3553	15	AAQ74208 Expression vector
6	454.8	31.9	7073	18	TF8-5G9 CDR-grafte
7	454.8	31.9	7864	18	AA750963 TF8-5G9 CDR-grafte
8	454.8	31.9	11795	21	AAA59345 Nucleotide sequenc
9	454.8	31.9	13254	17	AA740915 Nucleotide sequenc
10	454.8	31.9	13254	21	AAA32151 pEel2 Combo BM 12
11	454.8	31.9	13254	21	AAA32165 Complete nucleotid

c	12	453.8	31.8	608	19	AAV21731	ColEI origin of re
c	13	453.8	31.8	5241	17	AAAT27307	pHL104 containing
c	14	453.8	31.8	5365	10	AAAN90646	Nucleotide sequenc
c	15	453.8	31.8	5413	10	AAAN90649	Nucleotide sequenc
c	16	453.8	31.8	5518	10	AAAN90647	Nucleotide sequenc
c	17	453.8	31.8	5566	10	AAAN90648	Nucleotide sequenc
c	18	453.8	31.8	5579	11	AAQ02032	pBG391 pHS 164.
c	19	453.8	31.8	6051	12	AAQ14934	pBG391 sequence.
c	20	453.8	31.8	6149	10	AAAN90645	Sequence of T4 lym
c	21	453.8	31.8	6151	10	AAAN90644	Sequence of T4 lym
c	22	453.8	31.8	6151	11	AAQ03006	Full length T4 CDN
c	23	453.8	31.8	6151	11	AAQ05608	Plasmid pBG381 inc
c	24	453.8	31.8	6165	13	AAQ020324	Sol.rhesus CD4 seq
c	25	453.8	31.8	6477	11	AAQ02030	plasmid-pinf 4-49.
c	26	453.8	31.8	7252	20	AAQ86928	Complete sequence
c	27	453.8	31.8	7377	11	AAQ03005	Full length T4 CDN
c	28	453.8	31.8	7377	11	AAQ05607	Plasmid p170-2 inc
c	29	453.8	31.8	8119	17	AAAT33477	Subtilisin N62b/G1
c	30	453.4	31.8	1547	19	AAV21732	Humanised vector p
c	31	453.4	31.8	1807	19	AAV21733	Humanised vector p
c	32	453.4	31.8	2308	19	AAV21734	Humanised vector p
c	33	453.4	31.8	7507	16	AAQ91272	Plasmid vector pad
c	34	453.2	31.8	13254	21	AAA31025	pEel2 Combo BM 12
c	35	453	31.8	2077	20	AAAT23771	Vector pAS1b DNA.
c	36	453	31.8	2462	21	AAA74638	Plasmid pSP72. Es
c	37	453	31.8	2577	14	AAQ36620	Expression plasmid
c	38	453	31.8	2704	20	AAQ06758	Vector pUC28 nucle
c	39	453	31.8	2927	17	AAAT29158	Plasmid pTRP. Syn
c	40	453	31.8	2927	17	AAAT31789	Plasmid pTRP. Syn
c	41	453	31.8	2939	16	AAQ87350	Plasmid BGINV. Sy
c	42	453	31.8	2939	19	AAV37293	pBGINV plasmid use
c	43	453	31.8	3003	11	AAQ05745	Plasmid P.L-mu-smc
c	44	453	31.8	3104	10	AAAN90296	DNA target sequenc
c	45	453	31.8	3130	19	AAV46334	Universal reporter

ALIGNMENTS

RESULT	1
AAV21724	AAV21724 standard; CDNA; 1425 BP.
ID	AAV21724 standard; CDNA; 1425 BP.
XX	
AC	AAV21724;
XX	
DF	17-AUG-1998 (first entry)
XX	Humanised vector pITL.
DE	Vector; vaccine; tumour; antigen; plasmid pITL; ds.
XX	
KW	Chimeric - Homo sapiens.
OS	Chimeric - Escherichia coli.
OS	Synthetic.
XX	
PH	Key Location/Qualifiers
FT	misc_feature 1..221
FT	/*tag= a
FT	/note= "stuffer sequence"
FT	22..481
FT	polyA_site
FT	/*tag= b
FT	/note= "combined splice and polyA sequences"
FT	495..701
FT	CDS
FT	/*tag= c
FT	/note= "SupF gene"
FT	712..1164
FT	misc_feature
FT	/*tag= d
FT	/note= "ColEI origin of replication"
FT	1177..1425
FT	promoter
FT	/*tag= e
FT	/note= "RANTES promoter"
XX	
PN	WO9806863-A1.

XX 19-FEB-1998.
 XX PD
 XX PF 14-AUG-1997; 97WO-US14306.
 XX PR 14-AUG-1996; 96US-0023931.
 XX PA (USSH) US DEPT HEALTH & HUMAN SERVICES.
 XX PI Nelson EL, Nelson PJ;
 XX DR WPI; 1998-159552/14.
 XX Humanised polynucleotide vectors - comprising human derived promoter
 PT and sequence acceptance site, used for the production of vaccines
 PT
 XX Claim 14; Page 29-30; 125pp; English.
 CC Plasmid pITL comprises a base vector for novel humanised
 CC polynucleotide vectors. Such vectors comprise a human-derived
 CC promoter or mammalian homologue which is functional in mammalian
 CC target tissue and cells and a sequence acceptance site (see
 CC AAV21735-36), which accepts cDNA products from RT-PCR cloning. They
 CC also contain minimal non-human components, such as a replication
 CC origin (see AAV21715) and selectable marker gene (see AAV21717-18) that
 CC are necessary for production of the vector, as well as human-derived
 CC splice and polyA sequences (see AAV21723). The novel vectors are
 CC used to express target antigens, especially tumour antigens. They
 CC are non-replicating in mammalian cells but are capable of extended
 CC stable expression of target sequences generating an immune response
 CC in immunised individuals. The vectors selectively elicit immune
 CC responses to the target sequences with little or no immune response
 CC to the other components of the vectors. The target antigens are
 CC expressed as intracellular polypeptides or peptides and, as such,
 CC are processed as self polypeptides or peptides and appropriately
 CC presented on antigen presenting cells.
 XX SQ
 Sequence 1425 BP; 325 A; 381 C; 404 G; 315 T; 0 other;

Query Match 100.0%; Score 1425; DB 19; Length 1425;
 Best Local Similarity 100.0%; Pred. No. 0;
 Matches 1425; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TGCATGCGCGGATCTTATCACTGATACGTTGGTGGACATATATGTTATACGTGA 60
 DB 1 Tgcatacgaggttccttctacactgataaagtgttgacataattatgtttatcagtga 60
 QY 61 TAAAGTCTCAAGCATGACAAAGTTGACGCGAATACAGTATCCGTCGCGGCGCTGSACT 120
 DB 61 taaagtgtcaagcatgacaaagtgcagcgaataacagtgtatcgtgcgcgcctggact 120
 QY 121 GTTGACGAGGTGCGGTAGACGGTCTGACGACACGCAAACTGGCGGAACGGTTGGGGGT 180
 DB 121 gttgaacaggttcggttagacggtctgcagcacgcaaacatggcgaacggttgagggt 180
 QY 181 GCAGCAGCGCGCTTACTGGCACTTCAGACACAGCGGCGGCTTAAAGGCCATATGG 240
 DB 181 gcagcagcgcgcttacttgcacttcaggaacagcgcgcccttaaggccatagg 240
 QY 241 TGAGTGGATGCCCTGACCCAGCGGGGATGGGGAGACTCTGTAGTCAGAGCCCGCGGC 300
 DB 241 tgaagtgtatgcttgaccacagcgaggatggggagacctgtatcagagcccccggc 300
 QY 301 AGCAGAGGCAATGCCGCTCTCCCTGCAGTGTAGTACTGCCGGGTGGGATCCC 360
 DB 301 agcacaggccaatgcccgtccttccctgcagtgcagtgcagtgactcccggtgggatccc 360
 QY 361 TGTGACCCCTCCCGAGTGCCTCTCCCTGGCCCTGGAAGTTGCCAGTCCCGACACGAC 420
 DB 361 tgtgacctctcccgagtgctctcctgcccctggaggttgccactccagtgcccaccagc 420
 QY 421 CTGTGCTTAATAAAATTAAGTTGTCATCATTTTGTCTGACTAGGTGTCTCTATAATA 480

DB 421 ctgtctctaataaaattaaagttgcatcttttctgactaggtgtctctataataatta 480
 QY 481 TAAGCTTGATATCGAATCTTTCTCAACATAACACTTTTACAGGGGCGCTCATTTGATAT 540
 DB 481 taagcttgatcgaattcttctcaacgtaaacactttacagcgcgctcattgatatt 540
 QY 541 GATGCGCCCGCTTCCGATAAGGGAGCAGCGCCAGTAAAGCAATTTACCCGTGTGGGGTT 600
 DB 541 gatgcccccgcttcccgataaaggagcagggccagtaaaagcattaccctggtggggtt 600
 QY 601 CCGAGCGGCCAAAGGAGCAGACTCTATAATCTGCCGTATCGACTTCGAAGGTTCGAAT 660
 DB 601 cccgagcgccaaaggagcagacttaaatctgcgctcatcagacttcgaaggttcgaat 660
 QY 661 CTTTCCCGCCACACCATCTACTTTCAAAGTCCGAAAGTCTCTGCAGCCCGGTGAGCCG 720
 DB 661 ctttccccaccaccatcactttcaaaagtcgaaagaattctcgcagccggtgagccg 720
 QY 721 TAGTTAGGCGCACCATCTCAAGAACTCTGTAGCACCGCCTACATACCTGCTCTGCTAATC 780
 DB 721 tagttaaggccaccactcaagaactctgtagcaccgctcacatacctcgtctctgtaact 780
 QY 781 CTGTTACCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 840
 DB 781 ctgttaccagtgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgct 840
 QY 841 CGATAGTTACCGATTAAGCGGAGCGGCTGCGGCTGAACGGGGGTTCTGTCACACAGCCC 900
 DB 841 cgatagttaccggataaaggcgagcggtcggtgacggtggttcgtgcacagagccc 900
 QY 901 AGCTTGAGCGAAGCACTTACACCGAACTGAGATACCTACAGCTGAGCATTTGAGAAAGC 960
 DB 901 agcttgagcgaaacgacttaaccgaaactgagatacctacagctgagcattgagaaagc 960
 QY 961 GCGACGCTTCCCGAAGGGAGAAAGCGGACAGGTATCCGGTAAAGCGGAGGTTCGGAACA 1020
 DB 961 ggcacgcttcccgaaggagaaaggcgagaggtatccggttaagcgaggttcggaaca 1020
 QY 1021 GGAGAGCGCACGAGGAGCTTCCAGGGGAAACGCTGTGATCTTTATAGTCTCTGCGGG 1080
 DB 1021 ggagagcgcacgagggagagcttcccagggggaacgctggtatcttagtctctcggg 1080
 QY 1081 TTTGCCCACTCTGACTTGTAGCTCGATTTTGTGATGCTCGTCAAGGGGCGGAGCCTA 1140
 DB 1081 tttcgcaactctgacttgacttgagcgtcgattttgtgctgctcagggggcgagacct 1140
 QY 1141 TGGAAACCGCCAGCAACGCGCGCGGGGATCCGGAGAGCTACTCTAGATGAGAGACA 1200
 DB 1141 tggaaaaacgccaagcaacgcccgggggagatccggagagctcaactctagatgagagca 1200
 QY 1201 GTGAGGAGAGACAGAGACTCGAATTTCCGGAGCTPATTTTCAGTTTCTTTTCCGTTTGT 1260
 DB 1201 gtgagggagagacagagactcgaaattccggagctatttcagtttcttctccgtttgt 1260
 QY 1261 GCAATTTCACTTATGATACCGGCCCAATCTTGGTGTCTATTTTGGAAACTCCCTTAGGG 1320
 DB 1261 gcaatttcacttatgataccgcccgaatgcttggtgtctattttgaaaactccccctt 1320
 QY 1321 GATGCCCCCTCAACTGGCCCTATAAAGGCCACGCTGAGCTGCAGAGGATTCCTGCAGAG 1380
 DB 1321 gatgccccctcaactggccctataaaggccagcctgagctgagaggtattcctgcagagg 1380
 QY 1381 ATCAAGACAGCAGCTGGACCTCGCACGCTCTCCACAGGTACC 1425
 DB 1381 atcaagacagcacgtggacctcgcacagcctctccccacaggtacc 1425

RESULT 2
 AAV21727
 ID AAV21727 standard; cDNA; 2125 BP.
 XX
 AC AAV21727;


```

XX 17-AUG-1998 (first entry)
XX Humanised vector pITL-hHER/neu.
XX
XX Vector; vaccine: tumour; antigen: plasmid pITL-hHER/neu;
XX human; HER-2/neu; C-erbB-2; breast cancer; ds.
XX
XX Chimeric - Homo sapiens.
XX Synthetic.
XX
XX Key Location/Qualifiers
XX CDS 13..921
XX /*tag= a
XX /product= human HER-2/neu
XX polyA_site 922..1181
XX /*tag= b
XX /note= "combined splice and polyA sequences"
XX CDS 1195..1401
XX /*tag= c
XX /note= "SupF gene"
XX misc_feature 1412..1864
XX /*tag= d
XX /note= "Cole1 origin of replication"
XX promoter 1877..2125
XX /*tag= e
XX /note= "RANTES promoter"
XX
XX WO9806863-A1.
XX
XX 19-FEB-1998.
XX
XX 14-AUG-1997; 97WO-US14306.
XX
XX 14-AUG-1996; 96US-0023931.
XX
XX (USSH ) US DEPT HEALTH & HUMAN SERVICES.
XX
XX Nelson EL, Nelson PJ;
XX
XX WPI; 1998-159552/14.
XX
XX Humanised polynucleotide vectors - comprising human derived promoter
XX and sequence acceptance site, used for the production of vaccines
XX
XX Example 7; Page 41-42; 125pp; English.
XX
XX Plasmid pITL-hHER2/neu comprises base vector pITL (see AAV21724) and
XX a human HER-2/neu nucleic sequence. pITL-hHER2/neu was used to
XX evaluate the toxicity of anti-tumour vaccination in rats, and in
XX phase I and phase II trials to evaluate polynucleotide vaccination
XX in advanced breast cancer. Novel humanised vectors, which can be
XX based on pITL, comprise a human-derived promoter or mammalian
XX homologue which is functional in mammalian target tissue and cells
XX and a sequence acceptance site which accepts cDNA products from
XX RT-PCR cloning. The vectors are non-replicating in mammalian cells
XX but are capable of extended stable expression of the target
XX sequence, generating an immune response in immunised individuals.
XX The vectors selectively elicit immune responses to the target
XX sequences with little or no immune response to the other components
XX of the vectors.
XX
XX Sequence 2125 BP; 449 A; 650 C; 598 G; 428 T; 0 other;
XX
XX Query Match 84.4%; Score 1203; DB 19; Length 2125;
XX Best Local Similarity 100.0%; Pred. No. 0;
XX Matches 1203; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
XX
XX 223 GCCTTAGGGCCATATGGTAGTGAGTGCCTTAGCCCCAGCGGGGATGGGGGAGACCTG 282
XX }
XX 922 gccttaagggccatgatggtgagtgagtccttgaccccccagcggggagacctg 981

```

```

QY 283 TAGTCAGAGCCCGCGGACAGCACAGGCAATGCGCTCTCCCTGCAGTACGTAGTGA 342
DB tagtcagagccccggcagcacagggccaatgcccgtctctccctcagtgagtagtga 1041
QY 343 CTGCCCGGGTGGATCCCTGTGACCCCTCCAGTGCCTCTCTCGGCCCTGGAATGGCC 402
DB ctgccccgggtgggatccctgtgacccctccccagtgccctcctcctggccctggaagtgc 1101
QY 403 ACTCCAGTGGCCACAGCCTTGCTCTAATAAAATAGTTGCATCATTTTGTCTGACTAG 462
DB actccagtgccccaccagccttgctctaataaaattaaagtgcacattttgtctgactag 1161
QY 463 GTGTCTCTATAATATTATAGCTTGATATGAATCTTTCTCAAGCTTAACACATTACAG 522
DB gtgtcctctataataattatagccttgatataagcttgatcgaattctctcaacgtaaacactttacag 1221
QY 523 CGCGCGCTCATTTGATATGATGCGCCCGCTTCCCGATAAGGGAGCAGGCCAGTAAAGC 582
DB cgcgcgctcattgatgatgatgccccgcttcccgataagggagcagggccagtaaaagc 1281
QY 583 ATTACCGTGTGGGGTTCGAGCGGCGCAAGAGGAGCAGACTCTAATCTGCCGTATC 642
DB attaccggtggtgggttcccgagcgccaaaggagcagactctaaatctgcgctc 1341
QY 643 GACTTCGAAGGTTTCAATCTTCCCTCCACACCATCACTTTCAAAAGTCCGAAAGATTC 702
DB gacttcgaaggttcgaatctctctcccccaccaccatcatttcaaaagtccegaagaaatc 1401
QY 703 CTGACGCCCGTGTAGCGGTAGTTAGGCCACCATCTCAAGAACTCTGTAGCACCCCTACA 762
DB ctgacgcccggtgtagcggttagtgagccaccacttcaagaactctgtagcaccgctaca 1461
QY 763 TACCTCGCTCTGCTAATCTCTTACAGTGGCTGCTGCCACTGGGGATAAGTCTGTGCTT 822
DB tacctcgctctgctaactctctgtaccagtggtctgcagtggtgcgataaagtcgtctt 1521
QY 823 ACCGGTTGGACTCAAGACGATAGTTACCGGATAAGGCGCAGCGTCTGGCTGAACGGG 882
DB accgggttggactcaagacgatagttaccggataaaggcgacggtcggtgcgaacggggg 1581
QY 883 GGTTCGTGCACACGCCCGAGCTTGGAGCGAACGACCTACACCGAACTGAGATACCTACAG 942
DB ggttcgtgcacacgcccagcttgagcgaaacgacctacacggaactgagatacctacag 1641
QY 943 CGTAGCATTTGAGAAAGCGCCACGCTTCCCGAAGGAGAAAGCGGACAGGTATCCGGTA 1002
DB cgtgagcattgagaaagcgccacgcttcccgaaaggagaaagcgacaggtatccggta 1701
QY 1003 AGCGCAGGGTCCGAACAGGAGAGCGCACGAGGAGCTTCCAGGGGGAACGCTGGTAT 1062
DB agcggcagggtcggaaacaggagagcgcaaggaggagcttccaggggggaaacgcctggtat 1761
QY 1063 CTTTATAGTCTCTCGGGTTTTCGCCACTCTGACTTGTAGCTGCGATTTTGTGATGCTCG 1122
DB ctttatagctctcggtgttcgccacctctgactgagctgcatattttgtgactgcg 1821
QY 1123 TCAGGGGGCGGAGCCCTATGGAACACCCAGCAACGCGCGCGGGGATCCGGAGAGCTC 1182
DB tcagggggcgagccctatggaaaaacgccaacgcgccggggggatccggagagctc 1881
QY 1183 ACTCTAGATGAGAGCAGTGTAGGGAGAGACAGACTCTCGAATTTCCGAGCTATTTTCAG 1242
DB actctagatgagagagcagtgagggagagacagactcgaatttccggagactatttcag 1941
QY 1243 TTTCTTTTCCGTTTGTGCAATTTCACTTATGATACCGGCAATGCTTGGTGTCTATTT 1302
DB tttctttccgtttgtgcaatttcaacttatgatacccgccaatgcttgggtgtctatt 2001
QY 1303 TGGAAACTCCCTTTAGGGGATGCCCTCACTGAGCCCTATTAAGGGCCAGCCCTGAGTGC 1362
DB tggaaactcccccttaggggatgccccctcaactgccccataaaggccagcctgagctgc 2061

```

QY 1363 AGAGATTCTCGCAGAGATCAAGACACAGCAGTGGACCTCCACAGAGCTCTCCACAGGT 1422
 Db 2062 agaggattcctgcagagatcaagacagcagctggacctgcagcctctccacaggt 2121
 QY 1423 ACC 1425
 Db 2122 acc 2124
 |||
 RESULT 3
 AAV21726
 ID AAV21726 standard; cDNA; 1911 BP.
 XX AC AAV21726;
 XX DT 17-AUG-1998 (first entry)
 XX DE Humanised vector pITL-GFP.
 XX KW Vector; vaccine; tumour; antigen; plasmid pITL-GFP;
 XX KW green fluorescent protein; ds.
 XX OS Chimeric - Homo sapiens.
 OS Chimeric - Escherichia coli.
 OS Chimeric - Aequorea victoria.
 OS Synthetic.
 XX FH Key
 FT CDS Location/Qualifiers
 2..719
 /*tag= a
 /product= green fluorescent protein
 720..967
 polyA_site
 /*tag= b
 /note= "combined splice and polyA sequences"
 981..1187
 /*tag= c
 /note= "SupF gene"
 1198..1650
 /*tag= d
 /note= "ColE1 origin of replication"
 1663..1911
 /*tag= e
 /note= "RANTES promoter"
 WO9806863-A1.
 XX 19-FEB-1998.
 XX 14-AUG-1997; 97WO-US14306.
 XX 14-AUG-1996; 96US-0023931.
 XX (USSH) US DEPT HEALTH & HUMAN SERVICES.
 XX Nelson EL, Nelson PJ;
 XX WPI; 1998-159552/14.
 XX Humanised polynucleotide vectors - comprising human derived promoter
 and sequence acceptance site, used for the production of vaccines
 XX Example 4; Page 33-34; 125pp; English.
 XX Plasmid pITL-GFP comprises base vector pITL (see AAV21724) and a
 CC humanised green fluorescent protein (GFP) reporter sequence
 CC (see AAV21725). Novel humanised vectors comprise a human-derived
 CC promoter or mammalian homologue which is functional in mammalian
 CC target tissue and cells and a sequence acceptance site which
 CC accepts cDNA products from RT-PCR cloning. pITL-GFP was used to
 CC examine the kinetics of expression of a reporter sequence from
 CC such vectors in an animal model (Fisher 344 rats). Toxicity from
 CC polynucleotide vaccination was examined. The novel vectors are
 CC used to express target antigens, especially tumour antigens. They

CC are non-replicating in mammalian cells but are capable of extended
 CC stable expression of target sequences generating an immune response
 CC in immunised individuals. The vectors selectively elicit immune
 CC responses to the target sequences with little or no immune response
 CC to the other components of the vectors. The target antigens are
 CC expressed as intracellular polypeptides or peptides and, as such,
 CC are processed as self polypeptides or peptides and appropriately
 CC presented on antigen presenting cells.
 XX
 SQ Sequence 1911 BP; 474 A; 514 C; 511 G; 412 T; 0 other;

Query Match 83.8%; Score 1194.4; DB 19; Length 1911;
 Best Local Similarity 99.9%; Pred. No. 0;
 Matches 1195; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 230 GGGCCATATGCTGAGTGGATGCTTGACCCAGCGGGGATGGGGAGACCTGTAGTCAG 289
 Db 716 gagccatatgtgagtgatgccttgacccagcgggagtgaggagacctgtagtcag 775
 QY 290 AGCCCGCGGCGACGACAGCCCAATGCCGTCCTCCCTCCACGTGACTGACTGCCCG 349
 Db 776 agccccgggcagcacagcgaatgccctccctccctcagtgagtagtgactgccg 835
 QY 350 GGTGGATCCCTGTGACCCCTCCCGTCTCTCTGCGCCCTGGAAGTTGCCACTCCAG 409
 Db 836 gttgggacctgtgacctcccccagtcctctcctcctcagtgagtgactgccag 895
 QY 410 TGCCACGACCGCTTGCTTAATAAATTAAAGTTGCATCAATTTTGTCTGACTAGGTGCCT 469
 Db 896 tgccaccagccttgctcctaataaaattgaagttgcatcattttgtctgactaggtgctc 955
 QY 470 CTATAATATATAAGCTTGATATCGAATTTCTTCTCAAGTAACTTTACACGCGCGG 529
 Db 956 ctataatattataagcttgatcgaattcttctcaacgtacactttacagcggcg 1015
 QY 530 TCATTTGATATGATGCGCCCGCTTCCGATAGGAGGAGCGCCAGTAAAGCATTACCC 589
 Db 1016 tcatttgatgatgcgccccgcttcccgataaggagcagccagataaagcattacc 1075
 QY 590 GTGCTGGGTTCGCGAGCGGCCAAAGGGAGGACACTCTAAATCTGCCGTCATCGACTTCG 649
 Db 1076 gtggtgggttcccgagcgccaaaggagcagactcctaaatctgcgctcatcgacttcg 1135
 QY 650 AAGGTTTCGAATCCTTCCCGCACCCACCATCACTTTCAAAAGTCGAAAGAAATTCCTCAGC 709
 Db 1136 aaggttcgaatccttccccaccacacatcacttcaaaagtcgaaagaaattccgtgagc 1195
 QY 710 CCGGTAGCCGCTAGTTAGCCACCACTTCAAGAACTCTGTAGCACCGCCTACATACCTCG 769
 Db 1196 ccgtgtagcgttagtgccaccacttcaagaactctgtagcacgcgtacatacctcg 1255
 QY 770 CHTCTGCTAATCCTGTACCAGTGGCTGCTGCCAGTGGCGATAGTCTGCTCTTACCGGT 829
 Db 1256 ctctgctaactcctgttaccagtggtgctgctgcagtggtgagataagtcgtgtcttaccgg 1315
 QY 830 TGGACTCAAGACGATAGTTACCGGATAGGCGCAGCGGTGCGGCTGAACGGGGGTTCGT 889
 Db 1316 tggactcaagacgatagttaccgataagcgcagcgttcgggtcgaaacgggggttcgt 1375
 QY 890 GCACACAGCCCGCTGGAGCGAAGCACTACACCGAACTGAGATACCTACACGCTGAGC 949
 Db 1376 gcacacagcccagcttgagcgaacacctacacgcagactgagatacctacacgtgagc 1435
 QY 950 ATTGAGAAGCCCGCAGCTTCCCGAAGGAGAGAAAGCGGACAGGTATCCGGTAACGGCA 1009
 Db 1436 attggagaagcggccacgcttcccgagggagaaagcggacaggtatcccggttaagcgca 1495
 QY 1010 GGGTCGGAACAGGAGCGGCGACGAGGAGCTTCCAGGGGAAACCCCTGGTATCTTTATA 1069
 Db 1496 gggtcggaacagagagcgcagcagggagcttccaggggggaaacccctggatcttata 1555
 QY 1070 GTCCGTGTCGGGTTCGCCACCTCTGACTTGACGCTCGAATTTTGTGATGCTCGTCAGGG 1129

|||||
Db 1556 gtccgtcggttttcgccacactctgacttgagcgtcgattttgtgatgctcgtagggg 1615
Qy 1130 GCGGAGCCTATGAAACCCAGACACGCGCGGGGATCCGGAGAGTCACTCTAG 1189
Db 1616 ggcggagccctatgaaacccagcaacgcgcggggatcccgagagctcactctag 1675
Qy 1190 ATGAGAGCAGTCAAGGAGAGACAGACACTCGAATTTCCGGAGCTATTTCACTTTCTT 1249
Db 1676 atgagagcagtgagagagacagactcgaatttcgagactatttcagtttctt 1735
Qy 1250 TTCGGTTTGTGCAATTTCACTTATGATACGGCCCAATGCTGTGTTGCTATTTGGAAC 1309
Db 1736 ttccgttttgcgaatttcacttatgataccgcgccaatgcttggttgcattttggaac 1795
Qy 1310 TCCCTTAGGGGATGCCCTCAACTGCCCTATATAAGGCCAGCTGAGTGCAGAGGAT 1369
Db 1796 tccctttaggggagtcgccctcaactgcccataaaggccagctgagctgcagagat 1855
Qy 1370 TCCTGACAGGATCAAGACAGCAGCTGGACCTCGCACAGCCTTCCACAGGTACC 1425
Db 1856 tctgcagagatcaagacagcagctgagacctgcacagcctctcccacaggtacc 1911

RESULT 4
AAV21732
ID AAV21732 standard; cDNA; 1547 BP.
AC AAV21732;
XX
XX 17-AUG-1998 (first entry)
XX Humanised vector pITL-A.
XX Vector; vaccine; tumour; antigen; plasmid pITL-A; ds.
XX Chimeric - Homo sapiens.
OS Chimeric - Escherichia coli.
OS Synthetic.
XX
XX
FH Key Location/Qualifiers
FT CDS 472..680
FT /*tag= c
FT /note= "SupF gene"
FT misc_feature 686..1292
FT /*tag= d
FT /note= "ColE1 origin of replication"
FT promoter 1311..1547
FT /*tag= e
FT /note= "RANTES promoter"
XX
XX WO9806863-A1.
XX
XX 19-FEB-1998.
XX
XX 14-AUG-1997; 97WO-US14306.
XX
XX 14-AUG-1996; 96US-0023931.
XX
XX (USSH) US DEPT HEALTH & HUMAN SERVICES.
XX
XX Nelson EL, Nelson PJ;
XX
XX WPI; 1998-159552/14.
XX
XX Humanised polynucleotide vectors - comprising human derived promoter
XX and sequence acceptance site, used for the production of vaccines
XX
XX Claim 14; Page 56-57; 125pp; English.
XX
XX Plasmid pITL-A comprises a base vector for novel humanised
XX polynucleotide vectors. The vector in Escherichia coli DH
XX 10-beta/p3 is deposited as ATCC 98401. Novel vectors comprise a

CC human-derived promoter or mammalian homologue which is functional
CC in mammalian target tissue and cells and a sequence acceptance site
CC (see AAV21735-36) which accepts cDNA products from RT-PCR cloning
CC They also contain minimal non-human components, such as a replication
CC origin (see AAV21715) and selectable marker gene (see AAV21717-18) that
CC are necessary for production of the vector. The novel vectors are
CC used to express target antigens, especially tumour antigens. They
CC are non-replicating in mammalian cells but are capable of extended
CC stable expression of target sequences generating an immune response
CC in immunised individuals. The vectors selectively elicit immune
CC responses to the target sequences with little or no immune response
CC to the other components of the vectors. The target antigens are
CC expressed as intracellular polypeptides or peptides and, as such,
CC are processed as self polypeptides or peptides and appropriately
CC presented on antigen presenting cells.
XX
SQ Sequence 1547 BP; 344 A; 408 C; 425 G; 370 T; 0 other;

Query Match 34.4%; Score 490.4; DB 19; Length 1547;
Best Local Similarity 62.8%; Pred. No. 9.4e-126;
Matches 984; Conservative 0; Mismatches 401; Indels 183; Gaps 7;
Qy 3 CCATGGCGCGGATTCTTTATCATTGATTAAGTTGGTGGACATATTATGTTTATCATGTGATA 62
Db 12 ccatggcgggattctttatcattgataaagttggtggacattattgtttatcagtgata 71
Qy 63 AAGTGTCAAGCATGACAAAAGTTGAGCGCAATACAGTATCGTCCGCGCCCTGGAGTGT 122
Db 72 aagttcaagcatgacaaagttgagcgcgaatacagtgatcgtgcggccctggactgt 131
Qy 123 TGAACGAGGTGCGGTAGACGGTCTGACGACACGCAAACTGGCGAAGCTTGGGGGTGC 182
Db 132 tgaacgaggtgcggtgagacggttcgacgacacgcaactgctggaacggttggggggtgc 191
Qy 183 AGCAGCGCGCGCTTTACTGGCATTTCAGGAACAAGCGCGCGCTTAAGGGCCATATGGTG 242
Db 192 agcagcggcgctttactggtcacttcaggaacagcggcgcccttaaggccatattggtg 251
Qy 243 ACTGGATGCTTGTACCCCGAGCGGGGATGGGGAGACCTGTAGTCAGAGCCCGCGGAG 302
Db 252 agtggatgcttgaccccgagcggggtggtggagacgttagtcagagcccccgggcag 311
Qy 303 CACAGGCCAATGCCGCTCTCCCTGCAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGA 362
Db 312 cacagggccaatgcccgtccttcccctgcag----- 341
Qy 363 TGACCCCTCCCGAGTGCCTCTCTGGCCCTGGAAGTTGCCACTCCAGTCCCGCAGCCT 422
Db 342 -gatgagtagtgagtgccctcctcctgcccctggaagtgcactccagtgccaccagcct 400
Qy 423 TGTCTTAATAAAATTAAGTTCATCATTTTGTCTGACTAGTGTCTCTATAAATTATATA 482
Db 401 tgtcctaataaataaagttgcatcattttgtgactaggtgctcctataataataata 460
Qy 483 AGCTTGATATCAATTTCTTCTACAGCTTAACACTTTACAGCGCGCGCTCATTTGATATGA 542
Db 461 agcttgatcatgaattcttcggacttttgaagtgatggtggtggggga-aggattcga 519
Qy 543 TGGCCCGCTTCCCGATAAGGGAGCAGCGCCAGTAAAGAGCATTACCCGTGGTGGGTGCC 602
Db 520 accttcgaagtcgatgacgagcagatttagatctgctccctttggcgctcggaacccc 579
Qy 603 CGAGCGCCAAAGGAGCAGACTCTAAATCTGCCGTATCGACTTCGAAAGTTGCAATC- 661
Db 580 accacgggtaagtcttttactggtcgtcctccttactgggaagcggggtgcgcatcataca 639
Qy 662 CTTCCCGCCACCATCATCTTCAAAAGTCCGAAAGAAATTTCTTCGAGC-CCGTGTAGC 718
Db 640 aatgacgcgcgctgtaaaagtgttacgtttgagaagaattcctgcagcccgccggtgc 699
Qy 719 CGTAGTTAGGCCACCATCTCAAGAACTCTGTAG----- 752

Db 700 tggcgtttttccataggtctcgccccctgacgagcatcacaaaaatcgacgtcctaagtc 759
QY 753 ACCGCCCTACATACCTGCTCTGCTTAATCTGTACAGTGGCTGCTGCCAGTGGCGATAA 812
Db 760 agaggtggcgaaccgcagagactaaagataaccaaggcgtttcccccctggagctccc 819
QY 813 GTCTGTCTTACCGGGTGTGACCTACAGAGATATACCGGATAAGCGCGACGGTGGG 872
Db 820 tctggtcgtctcctgttccgaccctgcgttaccggataccctgcgcctttctccctt 879
QY 873 CTGACGGGGGGTCTGTCACACAGCCAGCTTGGAGCGAAGCACTACACCAACTGAG 932
Db 880 cgggaagcgtggcgctttctcaatgtcagctgtaggtatctcagttcgtgtadgtcg 939
QY 933 ATACCTACACGCTGAGCATTGAGAAAGCGCCAGCTTCC----- 971
Db 940 ttcgctccaagctggcgtgtgtacagaaacccccctgtagcccgagcgcgtgcgccttat 999
QY 972 -----CGAAGGGAGAAAGCGGACAGGTATCCGGTAAGCGGCAG 1010
Db 1000 ccggttaactatcgttctgagtcacaccccgtaagacacagacttatcgccactggcagcag 1059
QY 1011 GGTCCGACAGGAGCGGCACAGGAGCTTCCAGGGGGAACCCCTGGTATCTTTATAG 1070
Db 1060 ccactggttaacaggtatgacagcgcgaggtatgtadgcgtgtctacagattcttgaagt 1119
QY 1071 TCCTGTCCGGTTTCGCCACCTCTGACTTGAGCGTCGATTTTGTGATGTCGTCGACGGGG 1130
Db 1120 ggtggcctaactacgcctacactagaaggacagtatttggatctcgcctctgctgaagc 1179
QY 1131 CGGAGCGATNGAAAA----- 1148
Db 1180 cagttacccttcggaagaagagttggtagctctgtatcccggaacaaacaccgcgtggtta 1239
QY 1149 -----CGCCACCAACGGCGCGGGGAT 1171
Db 1240 gcggtggtttttgtttgcaagcagcagattacgcgcagaaaaaaggatctggtgggat 1299
QY 1172 CCGGAGAGCTCACTCTAGATGAGAGAGCAGTGGAGAGACAGACTCGAATTTCCGG 1231
Db 1300 ccggagagctcactctagatgagagcagtgagagagacagactcgaatttccgg 1359
QY 1232 AGCTATTTCAGTTTCTTTCCGTTTGTGCAATTTCACHTTATGATACCGGGCCAAATGCTT 1291
Db 1360 agctatttcagtttcttcttcctgtttgtgcaatttcaacttatgatcacogggccaatgctt 1419
QY 1292 GFTTGTCTATTTCGAAACTCCCTTAGGGGATGCCCTCAACTGGCCCTATATAAGGCCCA 1351
Db 1420 ggttgcatttttggaactcccccttggggatgccccctcaactggccctataaaggcca 1479
QY 1352 GCCTGAGCTGCAGAGGATTCTCTCAGAGGATCAAGACAGCACGTGGACCTCGCACAGCCT 1411
Db 1480 gctgagctgcagagattcctctcagaggtatcaagacagcagctggaacctcgcaagcct 1539
QY 1412 CTCCCCACA 1419
Db 1540 cteccaca 1547

RESULT 5
AAQ/4208/c
ID AAQ/4208 standard; cDNA; 3553 BP.
XX
AC AAQ/4208;
XX
DT 31-MAY-1995 (first entry)
DE Expression vector VJneo.
XX Influenza virus; antigen; specific; immune response; nucleoprotein;
KW hemagglutinin; polymerase; matrix protein; non-structural protein;
KW human; vaccine; ds.
XX

OS Synthetic.
XX WO9421797-A.
XX 29-SEP-1994.
XX 14-MAR-1994; 94WO-US02751.
XX 18-MAR-1993; 93US-0032383.
PR 08-JUL-1993; 93US-0089985.
XX (MERI) MERCK & CO INC.
PA (VICA-) VICAL INC.
XX
PI Donnelly JJ, Dworki VJ, Liu MA, Montgomery DL, Parker SE;
PI Shiver JW, Ulmer JB;
XX WPI; 1994-317017/39.
XX Polynucleotide vaccine comprising influenza virus genes - for
vaccination against more than one strain of influenza virus
XX Claim 24; Page 105-07; 171pp; English.
XX The sequences given in AAQ74206-08 represent expression vectors which
were used for the expression of influenza virus genes. VJ contains
the contains the ampr gene and VJneo contains the kanR gene in the
same orientation. DNA constructs such as these, containing the
influenza virus genes are capable of inducing the expression of an
antigenic influenza virus gene product which induces a specific immune
response upon introduction of the DNA construct into animal tissue in
vivo and resultant uptake of the DNA construct by cells which express
the encoded influenza gene. The encoded influenza virus gene encodes
CC nucleoprotein, hemagglutinin, polymerase, matrix or non-structural
CC human influenza virus gene products. The virus gene is operably linked
to one or more control sequences for incorporation in a vaccine.
XX Sequence 3553 BP; 869 A; 911 C; 839 G; 934 T; 0 other;
SQ

Query Match 32.3%; Score 459.8; DB 15; Length 3553;
Best Local Similarity 97.5%; Pred. No. 3.7e-117;
Matches 467; Conservative 0; Mismatches 12; Indels 0; Gaps 0;
QY 686 AAAGTCCGAAAGAAATTCCTCGACGCCGCTGTAGCCGTAGTTAGGCCACCACTTCAAGAACT 745
Db 2422 AGATACCAATACTGCTTCCTTCAGCTAGTGTAGCCGTAGTTAGGCCACCACTTCAAGAACT 2363
QY 746 CTGTAGACCGGCTACATACCTCGCTCTGCTAATCTGTATTACAGTGGCTGTGCGAGTG 805
Db 2362 CTGTAGACCGGCTACATACCTCGCTCTGCTAATCTGTATTACAGTGGCTGTGCGAGTG 2303
QY 806 GCGATAAGTCTGTCTTTACCGGGTTGGACTCAAGACGATAGTTACCGGATAAGGCCGAGC 865
Db 2302 GCGATAAGTCTGTCTTTACCGGGTTGGACTCAAGACGATAGTTACCGGATAAGGCCGAGC 2243
QY 866 GGTCCGGCTGAACGGGGGTTTGTGTGCACACAGCCAGCTTGGAGCGAAGCACTACACCG 925
Db 2242 GGTCCGGCTGAACGGGGGTTTGTGTGCACACAGCCAGCTTGGAGCGAAGCACTACACCG 2183
QY 926 AACTGAGATACCTACAGCTGAGCATTGAGAAAGCGCCACGCTTCCCGAAGGAGGAAAGG 985
Db 2182 AACTGAGATACCTACAGCTGAGCATTGAGAAAGCGCCACGCTTCCCGAAGGAGGAAAGG 2123
QY 986 CGGACAGGTATCCGGTAAGCGGCGAGGTGCGAACAGAGCGCAGAGGAGCTTCCAG 1045
Db 2122 CGGACAGGTATCCGGTAAGCGGCGAGGTGCGAACAGAGCGCAGAGGAGCTTCCAG 2063
QY 1046 GGGGAAACCGCTGGTATCTTTATAGTCCTGTGCGGGTTTCGCCACCTCTGACTTGAGCGTC 1105
Db 2062 GGGGAAACCGCTGGTATCTTTATAGTCCTGTGCGGGTTTCGCCACCTCTGACTTGAGCGTC 2003
QY 1106 GATTTTGTGATGCTCGTCAGGGGGCGGAGCCTATGGA AAAACGACGACGCGGCC 1164

```
Db 2002 GATTTTGTGATGCTGCTCAGGGGGCGGAGCCCTATGGAAGAAACGCCAGCAACGCGGCC 1944
|||||
RESULT 6
AAT50962/c
ID AAT50962 standard; DNA; 7073 BP.
XX
XX AAT50962:
XX
XX 01-APR-1997 (first entry)
XX
XX TF8-5G9 CDR-grafted heavy chain gene in pE6TF8HCDR20.
DE
XX
XX Tissue factor; complementarity determining region; CDR;
KW humanised antibody; monoclonal antibody; Tf8-5G9; coagulation;
KW thrombosis; restenosis; arteriosclerosis; therapy;
KW pE6TF8HCDR20; vector; ss.
XX
XX Chimeric Mus sp.;
OS Chimeric Homo sapiens;
OS Chimeric Escherichia coli;
OS Chimeric simian virus 40;
OS Chimeric human cytomegalovirus strain AD 169.
XX
XX Key Location/Qualifiers
FH misc_RNA 1..2360
FT /*tag= a
FT /function= TF8HCDR20 CDR-grafted heavy chain gene
FT misc_signal 7..15
FT /*tag= b
FT /function= Kozak sequence
FT sig_peptide 16..72
FT /*tag= c
FT misc_RNA 73..423
FT /*tag= d
FT /note= "CDR-grafted variable region"
FT misc_RNA 424..717
FT /*tag= e
FT /note= "human IgG4 CH1 domain"
FT intron 718..1110
FT /*tag= f
FT /note= "human IgG4 intron 2"
FT misc_RNA 1111..1146
FT /*tag= g
FT /note= "human IgG4 hinge"
FT intron 1147..1267
FT /*tag= h
FT /note= "human IgG4 intron 3"
FT misc_RNA 1268..1594
FT /*tag= i
FT /note= "human IgG4 CH2 domain"
FT intron 1268..1594
FT /*tag= j
FT /note= "human IgG4 intron 4"
FT misc_RNA 1692..2012
FT /*tag= k
FT /note= "human IgG4 CH3 domain"
FT 3'UTR 2013..2354
FT /*tag= l
FT /note= 1
FT polyA_signal 2770..2537
FT /*tag= m
FT misc_RNA 2594..3848
FT /*tag= n
FT /note= "fragment of pBR328 (nt 375-650, 1426-2422)
FT contains ColE1 replication origin"
FT misc_RNA 3849..4327
FT /*tag= o
FT /note= "BgII-XmnI fragment of pSP64 beta-lactamase
FT provides ampicillin resistance to
FT transformed bacteria"
FT misc_RNA 4328..4885
FT /*tag= p
```

```
FT misc_RNA /note= "XmnI-HindIII fragment of ColE1-based
FT plasmid pCT54"
FT 4886..7022
FT /*tag= g
FT /note= "Pst-I'm fragment of hCMV contg. middle
FT intermediate early promoter"
FT 7023..7073
FT /*tag= r
FT /note= "pSP64 polylinker"
XX
XX WO9640921-A1.
XX
XX 19-DEC-1996.
XX
XX 06-JUN-1996; 96WO-US09287.
XX
XX 07-JUN-1995; 95US-0480120.
XX
XX (JOHJ ) JOHNSON & JOHNSON.
XX
XX Jolliffe LK, Pulito VL, Zivin RA;
XX
XX WPI; 1997-099935/09.
XX P-PSDB; AAW10232.
XX
XX CDR-grafted antibody capable of inhibiting human tissue factor -
XX used in the treatment of thrombotic disorders, e.g.intravascular
XX coagulation, arterial restenosis etc.
XX
XX Claim 21; Page 74-80; 142pp; English.
XX
XX Vector pE6TF8HCDR20 (AAT50962) comprises the TF8HCDR20 CDR-grafted
XX heavy chain gene inserted into vector pE6hCMV-BgII. It encodes
XX a CDR-grafted heavy chain (AAW10232) in which CDRs (see also AAW10222-
XX 24) from the mouse anti-tissue factor monoclonal antibody Tf8-5G9
XX are grafted into a human KOL antibody framework. The heavy chain
XX constant region is derived from human IgG4. Humanised antibodies
XX are obt'd. by cotransfecting a host cell with vectors encoding the
XX CDR-grafted heavy and light chains (see also AAW10233). These
XX antibodies are capable of inhibiting human tissue factor and can be
XX used to treat e.g. intravascular coagulation, arterial restenosis
XX and arteriosclerosis.
XX
XX Sequence 7073 BP; 1673 A; 1925 C; 1778 G; 1697 T; 0 other;
```

```
Query Match 31.9%; Score 454.8; DB 18; Length 7073;
Best Local Similarity 99.6%; Pred. No. 1.2e-115;
Matches 456; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
QY 712 GTGTAGCCGTAGTTAGGCCACCACCTTCAAGAACTCTGTAGCAGCGCTACATACCTCGCT 771
Db |||||||
QY 772 CTGCTAATCTGTTTACCAGTGGCTGCTGCCAGTGGCGGATAAGTCTGTTTACCGGGTTG 831
Db |||||||
QY 3269 CTGCTAATCTGTTTACCAGTGGCTGCTGCCAGTGGCGGATAAGTCTGTTTACCGGGTTG 3210
QY 832 GACTCAAGACGATAGTTTACCGGATAAGCGCAGCGGTGGCGCTGAACGGGGGGTTCGTGC 891
Db |||||||
QY 3209 GACTCAAGACGATAGTTTACCGGATAAGCGCAGCGGTGGCGCTGAACGGGGGGTTCGTGC 3150
QY 892 ACACAGCCCGAGCTGGAGCGAAGCAGCTACACCGAACTGAGTACCTACAGCGTGACAT 951
Db |||||||
QY 3149 ACACAGCCCGAGCTGGAGCGAAGCAGCTACACCGAACTGAGTACCTACAGCGTGACAT 3090
QY 952 TGAGAAAGCCCGAGCTTCCCGAAGGAGAGAAAGCGGACAGGTATCCGGTAAAGCGCAGG 1011
Db |||||||
QY 3089 TGAGAAAGCCCGAGCTTCCCGAAGGAGAGAAAGCGGACAGGTATCCGGTAAAGCGCAGG 3030
QY 1012 GTCGGAACAGGAGCGCACGAGGAGCTTCCAGGGGGAACGCGCTGTATCTTTATAGT 1071
Db |||||||
QY 3029 GTCGGAACAGGAGCGCACGAGGAGCTTCCAGGGGGAACGCGCTGTATCTTTATAGT 2970
```

QY 1072 CCGTGGGTTTCGCCACCTGACTTGGGCTCGATTTTGTGATGCTGCTCAGGGGG 1131
 Db 2969 CCGTGGGTTTCGCCACCTGACTTGGGCTCGATTTTGTGATGCTGCTCAGGGGG 2910
 QY 1132 CGGAGCCTATGGAACACCCAGCAACCGCGCGGGG 1169
 Db 2909 CGGAGCCTATGGAACACCCAGCAACCGCGCGGGG 2872

RESULT 7

AAT50963/C
 ID AAT50963 standard; DNA; 7864 BP.

XX AC AAT50963;

XX DT 01-APR-1997 (first entry)

XX DE TF8-5G9 CDR-grafted light chain gene in pEel12TF8LCDR3.

XX KW Tissue factor; complementarity determining region; CDR;
 humanised antibody; monoclonal antibody; TF8-5G9; coagulation;
 thrombosis; restenosis; arteriosclerosis; therapy;
 pEel12TF8LCDR3; vector; ss.

XX OS Chimeric - Mus sp.

XX OS Chimeric - Homo sapiens.

XX OS Chimeric - Escherichia coli.

XX OS Chimeric - Simian virus 40.

XX OS Chimeric - Human cytomegalovirus strain AD 169.

XX OS Chimeric - Cricetus cricetus.

XX FH Key Location/Qualifiers

FT misc_RNA

FT 1..759

FT /tag= a

FT /product= TF8LCDR3 CDR-grafted heavy chain

FT 6..8

FT /tag= b

FT /function= "Kozak sequence"

FT 9..68

FT /tag= c

FT 69..392

FT /tag= d

FT /product= CDR-grafted variable region

FT 393..710

FT /tag= e

FT /product= human kappa constant region

FT 711..753

FT /tag= f

FT 760..3284

FT /tag= g

FT /note= "includes polyA signal, ColE1 replication origin and ampicillin resistance genes"

FT 3285..5736

FT /tag= h

FT /note= "includes CHO glutamine synthetase cDNA under control of the SV40 early promoter" transformed bacteria"

FT 5737..7864

FT /tag= i

FT /note= "base 4324 is given as n in the specification"

FT 5737..7864

FT /tag= j

FT /note= "includes hCMV middle intermediate early promoter and pSP64 polylinker"

FT 5737..7864

FT /tag= k

FT 5737..7864

FT /tag= l

FT 5737..7864

FT 5737..7864

FT 5737..7864

FT 5737..7864

PR 07-JUN-1995; 95US-0480120.

XX PA (JOHJ) JOHNSON & JOHNSON.

XX PI Joliffe LK, Pulito VL, Zivin RA;

XX DR WPI: 1997-099935/09.

XX DR P-PSDB; AAW10233.

XX DR P-PSDB; AAW10233.

XX DR P-PSDB; AAW10233.

XX DR P-PSDB; AAW10233.

XX DR P-PSDB; AAW10233.

XX DR P-PSDB; AAW10233.

XX DR P-PSDB; AAW10233.

XX DR P-PSDB; AAW10233.

XX DR P-PSDB; AAW10233.

XX DR P-PSDB; AAW10233.

XX DR P-PSDB; AAW10233.

XX DR P-PSDB; AAW10233.

XX DR P-PSDB; AAW10233.

XX DR P-PSDB; AAW10233.

XX DR P-PSDB; AAW10233.

XX DR P-PSDB; AAW10233.

XX DR P-PSDB; AAW10233.

XX DR P-PSDB; AAW10233.

XX DR P-PSDB; AAW10233.

XX DR P-PSDB; AAW10233.

XX DR P-PSDB; AAW10233.

XX DR P-PSDB; AAW10233.

XX DR P-PSDB; AAW10233.

XX DR P-PSDB; AAW10233.

XX DR P-PSDB; AAW10233.

XX DR P-PSDB; AAW10233.

XX DR P-PSDB; AAW10233.

XX DR P-PSDB; AAW10233.

XX DR P-PSDB; AAW10233.

XX DR P-PSDB; AAW10233.

XX DR P-PSDB; AAW10233.

XX DR P-PSDB; AAW10233.

XX DR P-PSDB; AAW10233.

XX DR P-PSDB; AAW10233.

XX DR P-PSDB; AAW10233.

XX DR P-PSDB; AAW10233.

XX DR P-PSDB; AAW10233.

XX DR P-PSDB; AAW10233.

XX DR P-PSDB; AAW10233.

XX DR P-PSDB; AAW10233.

XX DR P-PSDB; AAW10233.

XX DR P-PSDB; AAW10233.

XX DR P-PSDB; AAW10233.

XX DR P-PSDB; AAW10233.

XX DR P-PSDB; AAW10233.

XX DR P-PSDB; AAW10233.

XX DR P-PSDB; AAW10233.

XX DR P-PSDB; AAW10233.

XX DR P-PSDB; AAW10233.

XX DR P-PSDB; AAW10233.

XX DR P-PSDB; AAW10233.

XX DR P-PSDB; AAW10233.

XX DR P-PSDB; AAW10233.

XX DR P-PSDB; AAW10233.

XX DR P-PSDB; AAW10233.

XX DR P-PSDB; AAW10233.

XX DR P-PSDB; AAW10233.

XX DR P-PSDB; AAW10233.

XX DR P-PSDB; AAW10233.

XX DR P-PSDB; AAW10233.

Query Match 31.9%; Score 454.8; DB 18; Length 7864;
 Best Local Similarity 99.6%; Pred. No. 1.2e-115;
 Matches 456; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
 QY 712 GTGTAGCCGTAGTTAGGCCACACTTCAAGAACTCTGTAGCACCGCTTACATACCTCGCT 771
 Db 1728 GTGTAGCCGTAGTTAGGCCACACTTCAAGAACTCTGTAGCACCGCTTACATACCTCGCT 1669
 QY 772 CTGCTAATCTCTTACCAGTGGCTGCTGCCAGTGGCGAATAAGTGTCTTACCAGGTTG 831
 Db 1668 CTGCTAATCTCTTACCAGTGGCTGCTGCCAGTGGCGAATAAGTGTCTTACCAGGTTG 1609
 QY 832 GACTCAAGACGATAGTTACCGGATAAGCGCGAGCGGTGCGGCTGAACGGGGGTTCTGTC 891
 Db 1608 GACTCAAGACGATAGTTACCGGATAAGCGCGAGCGGTGCGGCTGAACGGGGGTTCTGTC 1549
 QY 892 ACACAGCCAGCTTGGAGCGAAGCACTTACACCACTGAGATACCTACACGCTGAGCAT 951
 Db 1548 ACACAGCCAGCTTGGAGCGAAGCACTTACACCACTGAGATACCTACACGCTGAGCAT 1489
 QY 952 TGAGAAAGCGCCACGCTTCCGAAAGGAGAAAGCGGACAGGTATCCGTAAGCGGCGAG 1011
 Db 1488 TGAGAAAGCGCCACGCTTCCGAAAGGAGAAAGCGGACAGGTATCCGTAAGCGGCGAG 1429
 QY 1012 GTCCGACAGAGAGCGCAGAGGAGCTTCCAGGGGGAACGCTGTATCTTTATAGT 1071
 Db 1428 GTCCGACAGAGAGCGCAGAGGAGCTTCCAGGGGGAACGCTGTATCTTTATAGT 1369
 QY 1072 CCTGTCCGGTTTCGCCACCTCTGACTTGGAGTCCGATTTTGTGATGCTGCTCAGGGGG 1131
 Db 1368 CCTGTCCGGTTTCGCCACCTCTGACTTGGAGTCCGATTTTGTGATGCTGCTCAGGGGG 1309
 QY 1132 CGGAGCCTATGGAACACCCAGCAACCGCGCGGGG 1169
 Db 1308 CGGAGCCTATGGAACACCCAGCAACCGCGCGGGG 1271

RESULT 8

AAA59345/c

ID AAA59345 standard; DNA; 11795 BP.

XX AC

XX AC AAA59345;

XX AC

XX AC

XX AC

XX AC

XX AC

DT 07-NOV-2000 (first entry)
XX Nucleotide sequence of plasmid pNIV4801.
DE Fusion protein; Varicella Zoster Virus; VZV; gE protein; chicken pox;
KW shingles; NSI protein; ss.
XX Synthetic.
OS Varicella Zoster Virus.
XX WO200043527-A1.
XX 27-JUL-2000.
XX 17-JAN-2000; 2000WO-EP00352.
XX 20-JAN-1999; 99GB-0001254.
XX (SMIK) SMITHKLINE BEECHAM BIOLOGICALS.
XX Bollen A, Haumont M, Jacobs P, Jacquet A, Massaer MGF;
XX WPI; 2000-505841/45.
XX Fusion protein comprising Varicella Zoster Virus gE protein and a
PT different VZV protein useful for treating Varicella (chickenpox) or
PT Zoster (shingles) -
XX Disclosure; Page 27-34; 60pp; English.
XX The specification describes a fusion protein, which comprises
CC a Varicella Zoster Virus (VZV) gE protein or an immunologically
CC active fragment, fused to a different protein of VZV. The
CC fusion protein or nucleic acids encoding it can be used to
CC prevent or ameliorate Varicella (chicken pox) or Zoster (shingles)
CC infections. The present sequence represents the plasmid pNIV4801, a
CC vector which is used to express fusion proteins of the invention.
XX Sequence 11795 BP; 2917 A; 2760 C; 2957 G; 3161 T; 0 other;
SQ

Query Match 31.9%; Score 454.8; DB 21; Length 11795;
Best Local Similarity 99.6%; Pred. No. 1.4e-115;
Matches 456; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 712 GTGTAGCGGTAGTTAGGCCACCACTTCAAGAACTCTGTAGCACCGCCTACATACCTCGCT 771
DB 986 GTGTAGCGGTAGTTAGGCCACCACTTCAAGAACTCTGTAGCACCGCCTACATACCTCGCT 927
QY 772 CTGCTAATCTCTGTACCAGTGGCTGCTGCCAGTGGCGAATAAGTGTCTTACCGGGTTG 831
DB 926 CTGCTAATCTCTGTACCAGTGGCTGCTGCCAGTGGCGAATAAGTGTCTTACCGGGTTG 867
QY 832 GACTCAAGACGATAGTTACCGGATAAGCGCGAGCGGTGCGGCTGAACGGGGGGTTCTGTC 891
DB 866 GACTCAAGACGATAGTTACCGGATAAGCGCGAGCGGTGCGGCTGAACGGGGGGTTCTGTC 807
QY 892 ACACAGCCCACTTGGACGGAACACCTACACCGAATGAGATACCTACACCGTACCAT 951
DB 806 ACACAGCCCACTTGGACGGAACACCTACACCGAATGAGATACCTACACCGTACCAT 747
QY 952 TGAGAAAGCGCCAGCGTTCCTCCGAGGAGAGAAAGGCGGACAGGTATCCGGTAAGCGGCAGG 1011
DB 746 TGAGAAAGCGCCAGCGTTCCTCCGAGGAGAGAAAGGCGGACAGGTATCCGGTAAGCGGCAGG 687
QY 1012 GTCGGACAGAGAGCGCACGAGGAGCTTCCAGGGGAAACGCGCTGTATCTTTATAGT 1071
DB 686 GTCGGACAGAGAGCGCACGAGGAGCTTCCAGGGGAAACGCGCTGTATCTTTATAGT 627
QY 1072 CCTGTCCGGTTTCCCACTCTGACTTACGCGTCAATTTTGTATGCTCTCTCAGGGGGG 1131
DB 626 CCTGTCCGGTTTCCCACTCTGACTTACGCGTCAATTTTGTATGCTCTCTCAGGGGGG 567

QY 1132 CGGAGCCTATGAAACAGCCAGCAACGCGCGCGGGG 1169
DB 566 CGGAGCCTATGAAACAGCCAGCAACGCGCGCGGAGG 529

RESULT 9
AAT40915/C
ID AAT40915 standard; DNA; 13254 BP.
XX AC AAT40915;
XX 29-JAN-1997 (first entry)
DE Nucleotide sequence of pEel2 Combo BM 12.
XX Heavy chain; light chain; variable region; VH: monoclonal antibody;
KW MAb; HIV; human immunodeficiency virus; glycoprotein; gp120; clone;
KW virus infectivity assay; precursor gp160; immunocompetence; human;
KW anti-HIV antibody; detection; HIV infection; ds.
XX Synthetic.
OS
XX WO9602273-A1.
XX 01-FEB-1996.
XX 11-JUL-1995; 95WO-US08743.
XX 18-JUL-1994; 94US-0276852.
XX (SCRI) SCRIPPS RES INST.
XX Barbas CF, Burton DR, Lerner RA;
XX WPI; 1996-179601/18.
XX Monoclonal antibody binding to VL/V2 loop of HIV gp120 - used in
PT passive immuno:therapy and detection of HIV infection.
XX Example 3; Page 276-285; 366pp; English.

This sequence represents the nucleotide sequence of the pEel2 vector,
pEel2 Combo BM 12, which contains the b12 heavy and light chain genes.
The pEel2 vector has a human CMV promoter for expression of the light
chain, a polylinker to provide cloning sites, and a polyadenylation
signal for termination of transcription. The vector also contains the
GS selectable marker gene whose expression is controlled by an SV40
early promoter at the 5' end of the GS gene, an intron, and a
polyadenylation signal at the 3' end of the GS gene. A heavy chain
cassette comprising the HCMV promoter, enhancer elements, heavy chain
gene and polyadenylation signal were removed from the pEE6 vector and
inserted into the pEel2 vector to generate the combinatorial construct
containing both the b12 light and heavy chain genes. The vector pEel2
Combo BM 12 was used to transfect CHO cells and an antibody, b12, was
expressed. The resulting antibody has the capacity to reduce HIV
infectivity titre in an in vivo virus infectivity assay by 50 % at a
concentration of less than 700 ng of antibody/ml. The MAb may be used
for determining immunocompetence of a human anti-HIV antibody and in
the detection of HIV infection.

Sequence 13254 BP; 3206 A; 3559 C; 3251 G; 3238 T; 0 other;

Query Match 31.9%; Score 454.8; DB 17; Length 13254;
Best Local Similarity 99.6%; Pred. No. 1.5e-115;
Matches 456; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 712 GTGTAGCGGTAGTTAGGCCACCACTTCAAGAACTCTGTAGCACCGCCTACATACCTCGCT 771
DB 6449 GTGTAGCGGTAGTTAGGCCACCACTTCAAGAACTCTGTAGCACCGCCTACATACCTCGCT 6390
QY 772 CTGCTAATCTCTGTACCAGTGGCTGCTGCCAGTGGCGAATAAGTGTCTTACCGGGTTG 831

Db 6389 CTGCTAATCTGTTTACCAAGTGGCTGCTGCCAGTGGCGATAAGTCGTCTTACCGGGTTG 6330
QY 832 GACTCAAGACGATAGTTACCGGATAAAGCGCAGCGTGGGCTGAACGGGGGTTTCGTGC 891
Db 6329 GACTCAAGACGATAGTTACCGGATAAAGCGCAGCGTGGGCTGAACGGGGGTTTCGTGC 6270
QY 892 ACACAGCCAGCTTGGAGCGAAGCACTTACACCGAACTGAGATACCTACAGCGTGAGCAT 951
Db 6269 ACACAGCCAGCTTGGAGCGAAGCACTTACACCGAACTGAGATACCTACAGCGTGAGCAT 6210
QY 952 TGAGAAAGCGCCACCTTCCGAGGAGAGAAAGCGGACAGGTATCCGGTAAAGCGGACG 1011
Db 6209 TGAGAAAGCGCCACCTTCCGAGGAGAGAAAGCGGACAGGTATCCGGTAAAGCGGACG 6150
QY 1012 CTCGGAACAGGAGCGCACGAGGAGCTTCCAGGGGAAACGGCTGTATCTTTATAGT 1071
Db 6149 CTCGGAACAGGAGCGCACGAGGAGCTTCCAGGGGAAACGGCTGTATCTTTATAGT 6090
QY 1072 CTTGTCGGGTTTCGCCACCTCTGACTTGAAGCTGATTTTGTGATGCTCTCAGGGGGG 1131
Db 6089 CTTGTCGGGTTTCGCCACCTCTGACTTGAAGCTGATTTTGTGATGCTCTCAGGGGGG 6030
QY 1132 CGGAGCCTATGGAAGAAAGCGCAGCAACGCGCGCGGGG 1169
Db 6029 CGGAGCCTATGGAAGAAAGCGCAGCAACGCGCGCGGGG 5992

RESULT 10

AAA32151/c
ID AAA32151 standard; DNA; 13254 BP.

XX
AC AAA32151;
XX

DT 04-JUL-2000 (first entry)

DE pBe12 Combo BM 12 containing b12 heavy and light chains.

XX Antibody; anti-HIV monoclonal antibody; glycoprotein-120;
KW human immunodeficiency virus type 1; HIV-1; infectivity titre;
KW passive immunotherapy; reduce severity; HIV-induced disease;
KW immunocompetence; active immunisation; ss.

OS Synthetic.

XX AU9948754-A.

XX 17-FEB-2000.

XX 16-SEP-1999; 99AU-0048754.

XX 16-SEP-1999; 99AU-0048754.

PR (SCRI) SCRIPPS RES INST.

XX Burton DR, Barbas CF, Lerner RA;

XX WPI; 2000-246867/22.

XX Human neutralizing monoclonal antibodies to human immunodeficiency
PT virus (HIV) used for providing passive immunotherapy to HIV are
PT specific for glycoprotein-120 -

PS Example 4; Figure 29; 374pp; English.

XX This sequence represents a polynucleotide used in the preparation of the
CC antibodies of the invention. The invention relates to the production of
CC an anti-HIV (human immunodeficiency virus) glycoprotein (gp)-120
CC monoclonal antibody capable of reducing an HIV infectivity titre in an
CC in vitro virus infectivity assay by 50% at a concentration of less than
CC 70 ng/mL. The method for the production of the antibody comprises:
CC (a) providing a first polynucleotide encoding a heavy chain
CC immunoglobulin amino acid sequence (which does not comprise the sequence
CC represented by AAY98206) and a second polynucleotide encoding a light

CC chain immunoglobulin amino acid sequence;
CC (b) inserting the first and second polynucleotide sequences into a host
CC cell;
CC (c) maintaining the host cell in conditions which allow the amino acid
CC sequences encoded by the polynucleotides to be expressed in the host
CC cell; and
CC (d) isolating the antibody comprising the heavy and light chain
CC immunoglobulin amino acid sequences from the host cell.
CC The anti-HIV gp-120 monoclonal antibody is used for providing passive
CC immunotherapy to HIV in a human. They can be administered to high-risk
CC patients to reduce the likelihood and/or severity of HIV-induced disease
CC and to patients who are already HIV-infected. The antibodies are used
CC for neutralising field isolates which provides information about the
CC immunocompetence of an immune response in HIV patients, for detecting
CC HIV in a biological fluid or tissue sample e.g. by radioimmunoassay, for
CC producing anti-idiotypic antibodies which can be used for active
CC immunisation and to screen human monoclonal antibodies to identify those
CC with the same binding specificity and to monitor the course of HIV
CC disease therapy by measuring the changes in concentration of HIV
CC in the body or in body fluids by immunoassay. The anti-HIV gp-120
CC monoclonal antibodies are encoded by a human polynucleotide sequence and
CC when used in vivo for diagnosis and immunotherapy of HIV-induced disease
CC reduce the problems of significant host immune response to the
CC antibodies associated with monoclonal antibodies of xenogeneic or
CC chimeric derivation.

XX
SQ Sequence 13254 BP; 3206 A; 3561 C; 3249 G; 3237 T; 1 other;

Query Match 31.9%; Score 454.8; DB 21; Length 13254;
Best Local Similarity 99.6%; Pred. No. 1.5e-115;
Matches 456; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 712 GTGTAGCCGTAGTTAGGCCACCACTTCAAGAACTCTGTAGCACCGCTACATACCTCGCT 771
Db 6449 GTGTAGCCGTAGTTAGGCCACCACTTCAAGAACTCTGTAGCACCGCTACATACCTCGCT 6390

QY 772 CTGCTAATCTGTTTACCAAGTGGCTGCTGCCAGTGGCGATAAGTCGTCTTACCGGGTTG 831

Db 6389 CTGCTAATCTGTTTACCAAGTGGCTGCTGCCAGTGGCGATAAGTCGTCTTACCGGGTTG 6330

QY 832 GACTCAAGACGATAGTTACCGGATAAGCGCGCTGCGGGCTGAACGGGGGTTCTGTCG 891

Db 6329 GACTCAAGACGATAGTTACCGGATAAGCGCGCTGCGGGCTGAACGGGGGTTCTGTCG 6270

QY 892 ACACAGCCAGCTTGGAGCGAAGCACTTACACCGAACTGAGATACCTACAGCGTGAGCAT 951

Db 6269 ACACAGCCAGCTTGGAGCGAAGCACTTACACCGAACTGAGATACCTACAGCGTGAGCAT 6210

QY 952 TGAGAAAGCGCCACGCTTCCGGAAGGAGAAAGCGGCGAGGTATCCGGTAAAGCGGCAGG 1011

Db 6209 TGAGAAAGCGCCACGCTTCCGGAAGGAGAAAGCGGCGAGGTATCCGGTAAAGCGGCAGG 6150

QY 1012 GTCGGAACAGGAGAGCGCACGAGGGAGCTTCCAGGGGAAACGCCCTGTATCTTTATAGT 1071

Db 6149 GTCGGAACAGGAGAGCGCACGAGGGAGCTTCCAGGGGAAACGCCCTGTATCTTTATAGT 6090

QY 1072 CTTGTCGGGTTTCGCCACCTCTGACTTGAAGCTGCGATTTTGTGATGCTCTCAGGGGGG 1131

Db 6089 CTTGTCGGGTTTCGCCACCTCTGACTTGAAGCTGCGATTTTGTGATGCTCTCAGGGGGG 6030

QY 1132 CGGAGCCTATGGAAGAAAGCGCAGCAACGCGCGCGGGG 1169

Db 6029 CGGAGCCTATGGAAGAAAGCGCAGCAACGCGCGCGGGG 5992

RESULT 11

AAA32165
ID AAA32165 standard; DNA; 13254 BP.

XX
AC AAA32165;

DT 04-JUL-2000 (first entry)

CC provides plasmid replication and growth within permissive strains
 CC of Escherichia coli. The novel vectors are used to express target
 CC antigens, especially tumour antigens. They are non-replicating in
 CC mammalian cells but are capable of extended stable expression of
 CC target sequences generating an immune response in immunised
 CC individuals. The vectors selectively elicit immune responses to
 CC the target sequences with little or no immune response to the other
 CC components of the vectors. The target antigens are expressed as
 CC intracellular polypeptides or peptides and, as such, are processed
 CC as self polypeptides or peptides and appropriately presented on
 CC antigen presenting cells.
 XX
 SQ Sequence 608 BP; 130 A; 175 C; 160 G; 143 T; 0 other;

Query Match 31.8%; Score 453.8; DB 19; Length 608;
 Best Local Similarity 99.6%; Pred. No. 8.8e-116;
 Matches 455; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 712 GTGTAGCGGTAGTTAGGCCACCACTTCAAGAACTCTGTAGCACCGCTACATACCTCGCT 771
 DB 457 GTGTAGCGGTAGTTAGGCCACCACTTCAAGAACTCTGTAGCACCGCTACATACCTCGCT 398
 QY 772 CTGCTAATCTGTTACAGTGGCTGCTGCCAGTGGCGATAAGTCTGTCTTACCGGGTTG 831
 DB 397 CTGCTAATCTGTTACAGTGGCTGCTGCCAGTGGCGATAAGTCTGTCTTACCGGGTTG 338
 QY 832 GACTCAAGACGATAGTTACCGGATAAGGCGACGCTGCGGCTGAACGGGGGTTCTGTGC 891
 DB 337 GACTCAAGACGATAGTTACCGGATAAGGCGACGCTGCGGCTGAACGGGGGTTCTGTGC 278
 QY 892 ACACAGCCAGCTTGGAGGGAAGCACTACACCGAACTGAGATACCTACAGCGTGAGCAT 951
 DB 277 ACACAGCCAGCTTGGAGGGAAGCACTACACCGAACTGAGATACCTACAGCGTGAGCAT 218
 QY 952 TGAGAAAGCCACGCTTCCCGAAGGAGAAAGCGGACAGGTATCCGGTAACGGCGAGG 1011
 DB 217 TGAGAAAGCCACGCTTCCCGAAGGAGAAAGCGGACAGGTATCCGGTAACGGCGAGG 158
 QY 1012 GTCGGAACAGGAGCGACGAGGAGCTTCCAGGGGAAACCGCTGGTATCTTTATAGT 1071
 DB 157 GTCGGAACAGGAGCGACGAGGAGCTTCCAGGGGAAACCGCTGGTATCTTTATAGT 98
 QY 1072 CCTGTGCGGTTTCGCCACCTCTGACTTGAGCGTCGATTTTTGTGATGCTCGTCAGGGGG 1131
 DB 97 CCTGTGCGGTTTCGCCACCTCTGACTTGAGCGTCGATTTTTGTGATGCTCGTCAGGGGG 38
 QY 1132 CGGAGCCTATGGAAGAACGCCACACGCGGCGGGG 1168
 DB 37 CGGAGCCTATGGAAGAACGCCACACGCGGCGGG 1

RESULT 13
 AAT27307
 ID AAT27307 standard; DNA; 5241 BP.
 XX
 AC AAT27307;
 XX
 DT 07-AUG-1996 (first entry)
 XX
 DE pHL104 containing mutated viral RNA.
 XX
 KW Vaccine; RNA virus; influenza virus; promoter; gene expression;
 XX attenuation; pHL104; pHL926; ds; cyclic.
 XX
 OS Synthetic.
 XX
 PN EP704533-A1.
 XX
 PD 03-APR-1996.
 XX
 PF 30-SEP-1994; 94EP-0115505.
 XX

PR 30-SEP-1994; 94EP-0115505.
 XX
 PA (FARB) BAYER AG.
 XX
 PI Hobom G, Menke A, Neumann G;
 XX
 DR WPI; 1996-173036/18.
 XX
 PT Attenuated RNA virus with improved expression rate - useful as
 PT vaccine, e.g. against HIV, herpes virus, rhinovirus or
 PT cytomegalovirus
 XX
 PS Disclosure; Page 21-31; 38pp; English.
 XX

Plasmid pHL104 (AAT27307) is based on pHL926, which comprises hybrid
 CC CAT cDNA with flanking non-coding sequences from influenza viral RNA
 CC segments, inserted in antisense orientation between murine rDNA
 CC promoter and terminator sequences. The CAT reporter gene replaces
 CC a haemagglutinin coding sequencing, retaining viral 5' and 3'
 CC untranslated sequences, which cooperatively constitute the vRNA
 CC promoter structure. vRNA 3' end mutations in pHL926 were
 CC created by PCR. Mouse B82 L cells were transfected with mutated
 CC constructs and subsequently with helper influenza A/FPV/Bratislava
 CC virus. A single nucleotide exchange (G3A) abolished promoter activity.
 CC A double mutant (G3A, U8C) resulted in significant CAT activity, while
 CC a triple mutant (G3A, C5U, U8C) further enhanced CAT activity. Thus
 CC certain specific modifications result in viruses with greatly enhanced
 CC expression rates. This may be utilized in viral vaccine vectors.
 XX
 SQ Sequence 5241 BP; 1264 A; 1320 C; 1379 G; 1278 T; 0 other;

Query Match 31.8%; Score 453.8; DB 17; Length 5241;
 Best Local Similarity 99.6%; Pred. No. 2e-115;
 Matches 455; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 712 GTGTAGCGGTAGTTAGGCCACCACTTCAAGAACTCTGTAGCACCGCTACATACCTCGCT 771
 DB 1619 GTGTAGCGGTAGTTAGGCCACCACTTCAAGAACTCTGTAGCACCGCTACATACCTCGCT 1678
 QY 772 CTGCTAATCTGTTACAGTGGCTGCTGCCAGTGGCGATAAGTCTGTCTTACCGGGTTG 831
 DB 1679 CTGCTAATCTGTTACAGTGGCTGCTGCCAGTGGCGATAAGTCTGTCTTACCGGGTTG 1738
 QY 832 GACTCAAGACGATAGTTACCGGATAAGGCGACGCTGCGGCTGAACGGGGGTTCTGTGC 891
 DB 1739 GACTCAAGACGATAGTTACCGGATAAGGCGACGCTGCGGCTGAACGGGGGTTCTGTGC 1798
 QY 892 ACACAGCCAGCTTGGAGGGAAGCACTACACCGAACTGAGATACCTACAGCGTGAGCAT 951
 DB 1799 ACACAGCCAGCTTGGAGGGAAGCACTACACCGAACTGAGATACCTACAGCGTGAGCAT 1858
 QY 952 TGAGAAAGCCACGCTTCCCGAAGGAGAAACCGCTGGTATCTTTATAGT 1011
 DB 1859 TGAGAAAGCCACGCTTCCCGAAGGAGAAACCGCTGGTATCTTTATAGT 1918
 QY 1012 GTCGGAACAGGAGCGACGAGGAGCTTCCAGGGGAAACCGCTGGTATCTTTATAGT 1071
 DB 1919 GTCGGAACAGGAGCGACGAGGAGCTTCCAGGGGAAACCGCTGGTATCTTTATAGT 1978
 QY 1072 CCTGTGCGGTTTCGCCACCTCTGACTTGAGCGTCGATTTTTGTGATGCTCGTCAGGGGG 1131
 DB 1979 CCTGTGCGGTTTCGCCACCTCTGACTTGAGCGTCGATTTTTGTGATGCTCGTCAGGGGG 2038
 QY 1132 CGGAGCCTATGGAAGAACGCCACACGCGGCGGGG 1168
 DB 2039 CGGAGCCTATGGAAGAACGCCACACGCGGCGGGG 2075

RESULT 14
 AAN90646/C
 ID AAN90646 standard; DNA; 5365 BP.
 XX

AA90645;
 01-FEB-1991 (first entry)
 Nucleotide sequence of region encoding first 113 AA of plasmid pBG394 including soluble T4-like (st4) polypeptide number 9 (st4#9).
 HIV; soluble T4-like polypeptide 9; immunotherapeutic; prophylactic; plasmid pBG394; diagnostic.
 Homo sapiens.
 W08901940-A.
 09-MAR-1989.
 01-SEP-1988; 88WO-US02940.
 07-JAN-1988; 88US-0141649.
 (BIOJ) BIOGEN INC.
 Fisher RA, Gilbert W, Sato VL, Flavell RA, Maraganore JM, Liu TR;
 WPI; 1989-085519/11.
 DNA sequences coding for soluble T4-like polypeptide(s) - used in immuno:therapeutic and immunosuppressive compsns. and for preventing, treating or detecting AIDS
 Fig 19; ; 207pp; English.
 It is the nucleotide sequence of the plasmid pBG394. The sequence was isolated from 2 libraries: a lambda gt cDNA library derived from T cell tumour line REX and a lambda gt10 cDNA library derived from peripheral blood lymphocytes (PBL). For screening, a series of chemically synthesised antisense oligonucleotide DNA probes based on the known T4 protein sequence was used. The polypeptide encoded is useful in immunotherapeutic, prophylactic and diagnostic compsns. It can be used to purify HIV from a sample.
 Sequence 5365 BP; 1300 A; 1380 C; 1355 G; 1330 T; 0 other;
 Query Match 31.8%; Score 453.8; DB 10; Length 5365;
 Best Local Similarity 99.6%; Pred. No. 2e-115;
 Matches 455; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
 Qy 712 GTGTAGCGGTAGTTAGGCGCCACTTCAAGAACTCTGTAGCAGCGCTTACATACCTCGCT 771
 Db 3975 GTGTAGCGGTAGTTAGGCGCCACTTCAAGAACTCTGTAGCAGCGCTTACATACCTCGCT 3916
 Qy 772 CTGCTAATCTCTTACCAGTGGCTGCTGCCAGTGGCGATAAGTCTGTCTTACCGGGTTG 831
 Db 3915 CTGCTAATCTCTTACCAGTGGCTGCTGCCAGTGGCGATAAGTCTGTCTTACCGGGTTG 3856
 Qy 832 GACTCAAGACGATAGTTACCGGTAAGCGCGAGCGGTGCGGTGAACGGGGTTCTGTC 891
 Db 3855 GACTCAAGACGATAGTTACCGGTAAGCGCGAGCGGTGCGGTGAACGGGGTTCTGTC 3796
 Qy 892 ACACAGCCCGAGCTTGGAGCGAACGACCTTACACCGAACTGAGATACCTACAGCGTGAGCAT 951
 Db 3795 ACACAGCCCGAGCTTGGAGCGAACGACCTTACACCGAACTGAGATACCTACAGCGTGAGCAT 3736
 Qy 952 TGAGAAAGCGCCAGCTTCCCGAAGGGAGAAAGGCGGACAGGTATCCGGTAAGCGGCAGG 1011
 Db 3735 TGAGAAAGCGCCAGCTTCCCGAAGGGAGAAAGGCGGACAGGTATCCGGTAAGCGGCAGG 3676
 Qy 1012 GTCGGAACAGAGAGCGCACAGGAGGAGCTTCCAGGGGAAACGCTGTATCTTTATAGT 1071
 Db 3675 GTCGGAACAGAGAGCGCACAGGAGGAGCTTCCAGGGGAAACGCTGTATCTTTATAGT 3616
 Qy 1072 CCTGTGCGGGTTTCCGCCACCTCTGACTTGAGCGTCTGATTTTGTGTGATGCTCTGTCAGGGGG 1131

Db 3615 CCTGTGCGGGTTTCCGCCACCTCTGACTTGAGCGTCTGATTTTGTGATGCTCTGTCAGGGGG 3556
 Qy 1132 CGGAGCCTATGCGAAACGCCAGCAACGCCGCGGGG 1168
 Db 3555 CGGAGCCTATGCGAAACGCCAGCAACGCCGCGGGGAG 3519
 RESULT 15
 AA90649/C
 ID AA90649 standard; DNA; 5413 BP.
 AC AA90649;
 XX
 DT 01-AUG-1990 (first entry)
 XX
 DE Nucleotide sequence of region encoding first 131 AA of plasmid pBG395 including soluble T4-like (st4) polypeptide number 10 (st4#10).
 XX
 DE HIV; soluble T4-like polypeptide 10; immunotherapeutic; prophylactic; plasmid pBG395; diagnostic.
 CC
 KW Homo sapiens.
 XX
 OS W08901940-A.
 XX
 PN 09-MAR-1989.
 XX
 PD 01-SEP-1988; 88WO-US02940.
 XX
 PF 07-JAN-1988; 88US-0141649, US-094322.
 XX
 PR (BIOJ) BIOGEN INC.
 XX
 PA Fisher RA, Gilbert W, Sato VL, Flavell RA, Maraganore JM, Liu TR;
 XX
 PI WPI; 1989-085519/11.
 XX
 DR DNA sequences coding for soluble T4-like polypeptide(s) - used in immuno:therapeutic and immunosuppressive compsns. and for preventing, treating or detecting AIDS
 XX
 PT Disclosure; ; 207pp; English.
 XX
 PS It is the nucleotide sequence of the plasmid pBG395. The sequence was isolated from 2 libraries: a lambda gt cDNA library derived from T cell tumour line REX and a lambda gt10 cDNA library derived from peripheral blood lymphocytes (PBL). For screening, a series of chemically synthesised antisense oligonucleotide DNA probes based on the known T4 protein sequence was used. The polypeptide encoded is useful in immunotherapeutic, prophylactic and diagnostic compsns. It can be used to purify HIV from a sample.
 CC
 SQ Sequence 5413 BP; 1309 A; 1401 C; 1365 G; 1338 T; 0 other;
 Query Match 31.8%; Score 453.8; DB 10; Length 5413;
 Best Local Similarity 99.6%; Pred. No. 2e-115;
 Matches 455; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
 Qy 712 GTGTAGCGGTAGTTAGGCGCCACTTCAAGAACTCTGTAGCAGCGCTTACATACCTCGCT 771
 Db 4023 GTGTAGCGGTAGTTAGGCGCCACTTCAAGAACTCTGTAGCAGCGCTTACATACCTCGCT 3964
 Qy 772 CTGCTAATCTCTTACCAGTGGCTGCTGCCAGTGGCGATAAGTCTGTCTTACCGGGTTG 831
 Db 3963 CTGCTAATCTCTTACCAGTGGCTGCTGCCAGTGGCGATAAGTCTGTCTTACCGGGTTG 3904
 Qy 832 GACTCAAGACGATAGTTACCGGTAAGCGCGAGCGGTGCGGTGAACGGGGTTCTGTC 891
 Db 3903 GACTCAAGACGATAGTTACCGGTAAGCGCGAGCGGTGCGGTGAACGGGGTTCTGTC 3844
 Qy 892 ACACAGCCCGAGCTTGGAGCGAACGACCTTACACCGAACTGAGATACCTACAGCGTGAGCAT 951

|||||
Db 3843 ACACAGCCAGCTTGGAGCGAAGACGACTACACCGAACTGAGATACCTACAGCGTGAGCAT 3784
QY 952 TGAGAAAGCGCCACGCTTCCCGAAGGGAGAAAGCGGGACAGGTATCCGGTAAGCGGCAGG 1011
|||||
Db 3783 TGAGAAAGCGCCACGCTTCCCGAAGGGAGAAAGCGGGACAGGTATCCGGTAAGCGGCAGG 3724
QY 1012 GTCCGAACAGGAGAGCGCACGAGGAGCTTCCAGGGGAAACGCTGTATCTTTATAGT 1071
Db 3723 GTCCGAACAGGAGAGCGCACGAGGAGCTTCCAGGGGAAACGCTGTATCTTTATAGT 3664
QY 1072 CCTGTCGGGTTTCGCCACCTCTGACTTGAGCGTCGATTTTGTGATGCTCGTCAGGGGG 1131
Db 3663 CCTGTCGGGTTTCGCCACCTCTGACTTGAGCGTCGATTTTGTGATGCTCGTCAGGGGG 3604
QY 1132 CGGAGCCTATGGAAAAACGCCAGCAACGCCGCCGGG 1168
|||||
Db 3603 CGGAGCCTATGGAAAAACGCCAGCAACGCCGCCGGG 3567

Search completed: January 17, 2002, 12:03:06
Job time: 17773 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: January 17, 2002, 11:52:06 ; Search time 221.34 Seconds
(without alignments)
1458.079 Million cell updates/sec

Title: US-09-242-202A-16

Perfect score: 1425

Sequence: 1 TGCCATGGCGCGGATTCCTTT.....CAGCCTCTCCACAGGTACC 1425

Scoring table: IDENTITY NUC

Gapop 10.0 , Gapext 1.0

Searched: 351203 seqs, 113238999 residues

Total number of hits satisfying chosen parameters: 702406

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

Issued_Patents_NA: *
1: /cgn2_6/ptodata/2/ina/5A_COMB.seq: *
2: /cgn2_6/ptodata/2/ina/5B_COMB.seq: *
3: /cgn2_6/ptodata/2/ina/6A_COMB.seq: *
4: /cgn2_6/ptodata/2/ina/6B_COMB.seq: *
5: /cgn2_6/ptodata/2/ina/PCTUS_COMB.seq: *
6: /cgn2_6/ptodata/2/ina/backfiles1.seq: *

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
C 1	454.8	31.9	13254	1	US-08-276-852-156
C 2	454.8	31.9	13254	1	US-08-276-852-170
C 3	454.8	31.9	13254	1	US-08-899-575-156
C 4	454.8	31.9	13254	1	US-08-899-575-170
C 5	454.8	31.9	13254	1	US-08-899-575-156
C 6	454.8	31.9	13254	1	US-08-899-575-170
C 7	454.8	31.9	13254	5	PCT-US95-08743-156
C 8	454.8	31.9	13254	5	PCT-US95-08743-170
C 9	453.8	31.8	1905	1	US-08-594-469-9
C 10	453.8	31.8	1905	2	US-08-906-957-9
C 11	453.8	31.8	4410	1	US-08-594-469-1
C 12	453.8	31.8	4410	2	US-08-906-957-1
C 13	453.8	31.8	6151	5	PCT-US91-02954-12
C 14	453.8	31.8	8119	1	US-08-460-343B-1
C 15	453.8	31.8	8119	1	US-08-398-028B-1
C 16	453.8	31.8	8119	2	US-08-504-265B-1
C 17	453	31.8	2927	2	US-08-941-647A-1
C 18	453	31.8	2939	1	US-08-119-512-2
C 19	453	31.8	2939	3	US-08-488-015B-2
C 20	453	31.8	2939	3	US-08-814-412-17
C 21	453	31.8	3003	6	5182260-18
C 22	453	31.8	3104	1	US-07-415-307A-1
C 23	453	31.8	3104	1	US-08-371-320-1
C 24	453	31.8	3130	4	US-09-038-141-1
C 25	453	31.8	3249	1	US-08-507-455-4
C 26	453	31.8	3301	2	US-08-447-430A-42
C 27	453	31.8	3400	1	US-08-507-455-3

28	453	31.8	3423	2	US-08-447-430A-40	Sequence 40, Appl
29	453	31.8	3474	2	US-08-447-430A-41	Sequence 41, Appl
30	453	31.8	3474	2	US-08-318-837-10	Sequence 10, Appl
31	453	31.8	3585	1	US-08-362-670B-9	Sequence 9, Appl
32	453	31.8	3585	3	US-08-333-576C-9	Sequence 9, Appl
33	453	31.8	3585	4	US-08-808-324-9	Sequence 9, Appl
34	453	31.8	3585	5	PCT-US94-14030A-9	Sequence 9, Appl
35	453	31.8	3623	2	US-07-989-847-13	Sequence 13, Appl
36	453	31.8	3623	4	US-08-469-411-13	Sequence 13, Appl
37	453	31.8	3632	1	US-07-745-382-13	Sequence 13, Appl
38	453	31.8	3632	1	US-07-921-848-13	Sequence 13, Appl
39	453	31.8	3632	1	US-08-115-680-3	Sequence 3, Appl
40	453	31.8	3632	1	US-07-941-372-3	Sequence 3, Appl
41	453	31.8	3632	1	US-08-165-301A-13	Sequence 13, Appl
42	453	31.8	3632	3	US-08-810-436-13	Sequence 13, Appl
43	453	31.8	3632	5	PCT-US93-08247-3	Sequence 3, Appl
44	453	31.8	3632	5	PCT-US94-14179-13	Sequence 13, Appl
45	453	31.8	3656	1	US-08-232-463-8	Sequence 8, Appl

ALIGNMENTS

RESULT 1
US-08-276-852-156/c
; Sequence 156, Application US/08276852
; Patent No. 5652138
; GENERAL INFORMATION:
; APPLICANT: Burton, Dennis R
; APPLICANT: Barbas, Carlos F
; APPLICANT: Lerner, Richard A
; TITLE OF INVENTION: HUMAN NEUTRALIZING MONOCLONAL ANTIBODIES
; TITLE OF INVENTION: TO HUMAN IMMUNODEFICIENCY VIRUS
; NUMBER OF SEQUENCES: 170
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: The Scripps Research Institute, Office of
; ADDRESS: 10666 No. 5652138th Torrey Pines Road, Suite 220,
; STREET: Mail Drop TPC8
; CITY: La Jolla
; STATE: CA
; COUNTRY: USA
; ZIP: 92037
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/276.852
; FILING DATE: 18-JUL-1994
; CLASSIFICATION: 514
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/178,302
; FILING DATE: 30-SEP-1993
; APPLICATION NUMBER: US 07/954,148
; FILING DATE: 30-SEP-1992
; ATTORNEY/AGENT INFORMATION:
; NAME: Fitting, Thomas
; REGISTRATION NUMBER: 34,163
; REFERENCE/DOCKET NUMBER: SCRI452P
; TELEPHONE: 619-554-2937
; TELEFAX: 619-554-6312
; INFORMATION FOR SEQ ID NO: 156:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 13254 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: circular
; MOLECULE TYPE: DNA (genomic)
US-08-276-852-156

```

Query Match          31.9%; Score 454.8; DB 1; Length 13254;
Best Local Similarity 99.6%; Pred. No. 9.5e-133;
Matches 456; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 712 GTGTAGCGTAGTTAGGCGACACATTCAAGAACTCTGTAGCACCGCCCTACATACCTCGCT 771
DB 6449 GTGTAGCGTAGTTAGGCGACACATTCAAGAACTCTGTAGCACCGCCCTACATACCTCGCT 6390
QY 772 CTGCTTAATCCCTGTTACCAAGTGGCTCTGCCAGTGGCGATAAGTCGTGCTTTACCGGGTTG 831
DB 6389 CTGCTTAATCCCTGTTACCAAGTGGCTCTGCCAGTGGCGATAAGTCGTGCTTTACCGGGTTG 6330
QY 832 GACTCAAGACGATAGTTACCGGATAAAGCGGAGCGGTGCGGTCCGGCTGACCGGGGGTTTCGTGC 6270
DB 6329 GACTCAAGACGATAGTTACCGGATAAAGCGGAGCGGTGCGGTCCGGCTGACCGGGGGTTTCGTGC 6270
QY 892 ACACAGCGCCAGTTGAGCGGAACGACCTACACCGAACTAGATACCTACACGCGTGAGCAT 951
DB 6269 ACACAGCGCCAGTTGAGCGGAACGACCTACACCGAACTAGATACCTACACGCGTGAGCAT 6210
QY 952 TGAGAAAGCGCCAGCTTCCCGAAGGGAGAAAGCGGACAGGTATCCGGTAAAGCGCGAGG 6150
DB 6209 TGAGAAAGCGCCAGCTTCCCGAAGGGAGAAAGCGGACAGGTATCCGGTAAAGCGCGAGG 6150
QY 1012 GTCGGAACAGGAGCGCACGAGGAGCTTCCAGGGGAAACGCGCTGGTATCTTTATAGT 1071
DB 6149 GTCGGAACAGGAGCGCACGAGGAGCTTCCAGGGGAAACGCGCTGGTATCTTTATAGT 6090
QY 1072 CCTGTCGGGTTTCGCCACCTCTGACTTGTAGCGTCGATTTTGTGATGCTGTCAGGGGGG 1131
DB 6089 CCTGTCGGGTTTCGCCACCTCTGACTTGTAGCGTCGATTTTGTGATGCTGTCAGGGGGG 6030
QY 1132 CGGAGCCTATGAAAAACGCCAGCAACGCGCGCGGGG 1169
DB 6029 CGGAGCCTATGAAAAACGCCAGCAACGCGCGCGGG 5992

RESULT 2
US-08-276-852-170
; Sequence 170, Application US/08276852
; Patent No. 5652138
; GENERAL INFORMATION:
; APPLICANT: Burton, Dennis R
; APPLICANT: Barbas, Carlos F
; APPLICANT: Lerner, Richard A
; TITLE OF INVENTION: HUMAN NEUTRALIZING MONOCLONAL ANTIBODIES
; TITLE OF INVENTION: TO HUMAN IMMUNODEFICIENCY VIRUS
; NUMBER OF SEQUENCES: 170
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: The Scripps Research Institute, Office of
; ADDRESSEE: Patent Counsel
; STREET: 10666 No. 5652138th Torrey Pines Road, Suite 220,
; STREET: Mail Drop TPC8
; CITY: La Jolla
; STATE: CA
; COUNTRY: USA
; ZIP: 92037
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/276,852
; FILING DATE: 18-JUL-1994
; CLASSIFICATION: 514
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/178,302
; FILING DATE: 30-SEP-1993
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/954,148

```

```

? COUNTRY: USA
? ZIP: 92037
? COMPUTER READABLE FORM:
? MEDIUM TYPE: Floppy disk
? COMPUTER: IBM PC compatible
? OPERATING SYSTEM: PC-DOS/MS-DOS
? SOFTWARE: PatentIn Release #1.0, Version #1.25
? CURRENT APPLICATION DATA:
? APPLICATION NUMBER: US/08/899,575
? FILING DATE: 24-JUL-1997
? CLASSIFICATION: 435
? PRIOR APPLICATION DATA:
? APPLICATION NUMBER: US 08/276,852
? FILING DATE: 18-JUL-1994
? APPLICATION NUMBER: US 08/178,302
? FILING DATE: 30-SEP-1993
? PRIOR APPLICATION DATA:
? APPLICATION NUMBER: US 07/954,148
? FILING DATE: 30-SEP-1992
? ATTORNEY/AGENT INFORMATION:
? NAME: Fitting, Thomas
? REGISTRATION NUMBER: 34,163
? REFERENCE/DOCKET NUMBER: SCR1452P
? TELECOMMUNICATION INFORMATION:
? TELEPHONE: 619-554-2937
? TELEFAX: 619-554-6312
? INFORMATION FOR SEQ ID NO: 156:
? SEQUENCE CHARACTERISTICS:
? LENGTH: 13254 base pairs
? TYPE: nucleic acid
? STRANDEDNESS: double
? TOPOLOGY: circular
? MOLECULE TYPE: DNA (genomic)
? US-08-899-575-156

```

Query Match	31.9%	Score 454.8	DB 1	Length 13254
Best Local Similarity	99.6%	Pred. No. 9.5e-133		
Matches 456	Conservative 0	Mismatches 2	Indels 0	Gaps 0
Qy 712	GTGTAGCGGTAGTTAGGCCACCACTTCAAGAACTCTGTAGCACCGCCTACATACCTCGCT	771		
Db 6449	GTGTAGCGGTAGTTAGGCCACCACTTCAAGAACTCTGTAGCACCGCCTACATACCTCGCT	6390		
Qy 772	CTGCTAATCCCTGTTACCAAGTGCGTCTGCCAGTGGCGATAAGTCGTGTCTTTACCGGGTGG	831		
Db 6389	CTGCTAATCCCTGTTACCAAGTGCGTCTGCCAGTGGCGATAAGTCGTGTCTTTACCGGGTGG	6330		
Qy 832	GACTCAAGACGATAGTTACCGGATAAGCGCAGCGGTCTGGGCTGAACGGGGGTTCTGTGC	891		
Db 6329	GACTCAAGACGATAGTTACCGGATAAGCGCAGCGGTCTGGGCTGAACGGGGGTTCTGTGC	6270		
Qy 892	ACACAGCCGACTTTGGAGCGAAGCACTACACGAACTGAGATACCTACACGCTGAGCAT	951		
Db 6269	ACACAGCCGACTTTGGAGCGAAGCACTACACGAACTGAGATACCTACACGCTGAGCAT	6210		
Qy 952	TGAGAAAGCGCCACGCTTCCGAAAGGGAGAAAGCGGACAGGTATCCGGTAAAGCGGCAGG	1011		
Db 6209	TGAGAAAGCGCCACGCTTCCGAAAGGGAGAAAGCGGACAGGTATCCGGTAAAGCGGCAGG	6150		
Qy 1012	GTCCGGAACAGAGAGCGCACGAGGAGCTTCCAGGGGGAAACGGCTGGTATCTTTATAGT	1071		
Db 6149	GTCCGGAACAGAGAGCGCACGAGGAGCTTCCAGGGGGAAACGGCTGGTATCTTTATAGT	6090		
Qy 1072	CCTCTCGGGTTTCCCAACCTCTGCATTTGAGCGTCGATTTTTGTGATGCTCGTCAGGGGGG	1131		
Db 6089	CCTCTCGGGTTTCCCAACCTCTGCATTTGAGCGTCGATTTTTGTGATGCTCGTCAGGGGGG	6030		
Qy 1132	CGAGCCTATGGAAAAACGCCAGCAACGCGCGCGGGG	1169		
Db 6029	CGGAGCCTATGGAAAAACGCCAGCAACGCGCGCGGAG	5992		

RESULT 4
 US-08-899-575-170
 : Sequence 170, Application US/08899575
 : Patent No. 5770440
 : GENERAL INFORMATION:
 : APPLICANT: Burton, Dennis R
 : APPLICANT: Barbas, Carlos F
 : APPLICANT: Lerner, Richard A
 : TITLE OF INVENTION: HUMAN NEUTRALIZING MONOCLONAL ANTIBODIES
 : TITLE OF INVENTION: TO HUMAN IMMUNODEFICIENCY VIRUS
 : NUMBER OF SEQUENCES: 170
 : CORRESPONDENCE ADDRESS:
 : ADDRESSEE: The Scripps Research Institute, Office of
 : ADDRESSEE: Patent Counsel
 : STREET: 10666 No. 5770440th Torrey Pines Road, Suite 220,
 : STREET: Mail drop TPC8
 : CITY: La Jolla
 : STATE: CA
 : COUNTRY: USA
 : ZIP: 92037
 : COMPUTER READABLE FORM:
 : MEDIUM TYPE: Floppy disk
 : COMPUTER: IBM PC compatible
 : OPERATING SYSTEM: PC-DOS/MS-DOS
 : SOFTWARE: PatentIn Release #1.0, Version #1.25
 : CURRENT APPLICATION DATA:
 : APPLICATION NUMBER: US/08/899,575
 : FILING DATE: 24-JUL-1997
 : CLASSIFICATION: 435
 : PRIOR APPLICATION DATA:
 : APPLICATION NUMBER: US 08/276,852
 : FILING DATE: 18-JUL-1994
 : APPLICATION NUMBER: US 08/178,302
 : FILING DATE: 30-SEP-1993
 : PRIOR APPLICATION DATA:
 : APPLICATION NUMBER: US 07/954,148
 : FILING DATE: 30-SEP-1992
 : ATTORNEY/AGENT INFORMATION:
 : NAME: Fitting, Thomas
 : REGISTRATION NUMBER: 34,163
 : REFERENCE/DOCKET NUMBER: SCRI452p
 : TELECOMMUNICATION INFORMATION:
 : TELEPHONE: 619-554-2937
 : TELEFAX: 619-554-6312
 : INFORMATION FOR SEQ ID NO: 170:
 : SEQUENCE CHARACTERISTICS:
 : LENGTH: 13254 base pairs
 : TYPE: nucleic acid
 : STRANDEDNESS: double
 : TOPOLOGY: circular
 : MOLECULE TYPE: DNA (genomic)
 : US-08-899-575-170

Query Match	31.9%;	Score	454.8;	DB	1;	Length	13254;
Best Local Similarity	99.6%;	Pred. No.	9.5e-133;				
Matches	456;	Conservative	0;	Mismatches	2;	Indels	0;
Gaps	0;						

Qy	712	GTGTAGCCGCTAGTTAGGCCACACATTCAAGAATCTGTAGCAGCCGCTACATACCTCGCT	771
Db	6806	GTGTAGCCGCTAGTTAGGCCACCATTCAAGAATCTGTAGCACCCGCTACATACCTCGCT	6865
Qy	772	CTGCTTAATCTGTTTACCAGTGCTGCCAGTGGCGATAAGTCGTGCTTTACCGGGTTG	831
Db	6866	CTGCTTAATCTGTTTACCAGTGCTGCCAGTGGCGATAAGTCGTGCTTTACCGGGTTG	6925
Qy	832	GACTCAACACCATAGTTTACCGSATAAGCGCAGCGGTGGGCTGAACCGGGGTTCTGC	891
Db	6936	GACTCAACACCATAGTTTACCGSATAAGCGCAGCGGTGGGCTGAACCGGGGTTCTGC	6985
Qy	892	ACACAGCCCGCTTGGAGCGAACGACCTTACACCGAACTGAGATACCTACAGCGTGAGCAT	951
Db	6986	ACACAGCCCGCTTGGAGCGAACGACCTTACACCGAACTGAGATACCTACAGCGTGAGCAT	7045

QY 952 TCAGAAAGCCACCTTCCCGAAGGAGAAAGCGGACAGGTATCCGGTAAGCGGAGG 1011
Db 7046 TGAGAAAGCCACCTTCCCGAAGGAGAAAGCGGACAGGTATCCGGTAAGCGGAGG 7105
QY 1012 GTCGGAACAGGAGCGCACGAGGAGCTTCCAGGGGAAACGCTGTATCTTTATAGT 1071
Db 7106 GTCGGAACAGGAGCGCACGAGGAGCTTCCAGGGGAAACGCTGTATCTTTATAGT 7165
QY 1072 CCGTCCGGGTTTCGCCACCTCTGACTTGAGCGTGCATTTTGTGATGCTCTCAGGGGG 1131
Db 7166 CCGTCCGGGTTTCGCCACCTCTGACTTGAGCGTGCATTTTGTGATGCTCTCAGGGGG 7225
QY 1132 CGGAGCCTATGAAAGCCAGCAACCGCGCGGGG 1169
Db 7226 CGGAGCCTATGAAAGCCAGCAACCGCGCGGGG 7263

RESULT 5

US-08-899-575-156/c
; Sequence 156, Application US/08899575
; Patent No. 5804440
; GENERAL INFORMATION:
; APPLICANT: Burton, Dennis R
; APPLICANT: Barbas, Carlos F
; APPLICANT: Lerner, Richard A
; TITLE OF INVENTION: HUMAN NEUTRALIZING MONOCLONAL ANTIBODIES
; TITLE OF INVENTION: TO HUMAN IMMUNODEFICIENCY VIRUS
; NUMBER OF SEQUENCES: 170
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: The Scripps Research Institute, Office of
; STREET: 10666 No. 580440th Torrey Pines Road, Suite 220,
; STREET: Mail Drop TPC8
; CITY: La Jolla
; STATE: CA
; COUNTRY: USA
; ZIP: 92037
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/899,575
; FILING DATE: 24-JUL-1997
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/276,852
; FILING DATE: 18-JUL-1994
; APPLICATION NUMBER: US 08/178,302
; FILING DATE: 30-SEP-1993
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/954,148
; FILING DATE: 30-SEP-1992
; ATTORNEY/AGENT INFORMATION:
; NAME: Fitting, Thomas
; REGISTRATION NUMBER: 34,163
; REFERENCE/DOCKET NUMBER: SCR1452P
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 619-554-2937
; TELEFAX: 619-554-6312
; INFORMATION FOR SEQ ID NO: 156:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 13254 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: circular
; MOLECULE TYPE: DNA (genomic)
US-08-899-575-156

Query Match 31.9%; Score 454.8; DB 1; Length 13254;

Best Local Similarity 99.6%; Pred. No. 9.5e-133;
Matches 456; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
QY 712 GTGTAGCCGTAGTTAGCCACCACTTCAAGAACTGTAGCAGCGCTACATACCTCGCT 771
Db 6449 GTGTAGCCGTAGTTAGCCACCACTTCAAGAACTGTAGCAGCGCTACATACCTCGCT 6390
QY 772 CTGCTAATCTCTGTACCAGTGGCTGCTGCCAGTGGCGATAAGTCTGTCTTACCGGGTTG 831
Db 6389 CTGCTAATCTCTGTACCAGTGGCTGCTGCCAGTGGCGATAAGTCTGTCTTACCGGGTTG 6330
QY 832 GACTCAAGACGATAGTTACCGGATAAAGCGGACGCGTTCGGGCTGAACGGGGGTTCTGTC 891
Db 6329 GACTCAAGACGATAGTTACCGGATAAAGCGGACGCGTTCGGGCTGAACGGGGGTTCTGTC 6270
QY 892 ACACAGCCAGCTTGGAGCGAAGCAGCTACACGAACTGAGATACCTACAGCGTGAGCAT 951
Db 6269 ACACAGCCAGCTTGGAGCGAAGCAGCTACACGAACTGAGATACCTACAGCGTGAGCAT 6210
QY 952 TGAGAAAGCCACCTTCCCGAAGGAGAAAGCGGACAGGTATCCGGTAAGCGGAGG 1011
Db 6209 TGAGAAAGCCACCTTCCCGAAGGAGAAAGCGGACAGGTATCCGGTAAGCGGAGG 6150
QY 1012 GTCGGAACAGGAGCGCACGAGGAGCTTCCAGGGGAAACGCTGTATCTTTATAGT 1071
Db 6149 GTCGGAACAGGAGCGCACGAGGAGCTTCCAGGGGAAACGCTGTATCTTTATAGT 6090
QY 1072 CCGTCCGGGTTTCGCCACCTCTGACTTGAGCGTGCATTTTGTGATGCTCTCAGGGGG 1131
Db 6089 CCGTCCGGGTTTCGCCACCTCTGACTTGAGCGTGCATTTTGTGATGCTCTCAGGGGG 6030
QY 1132 CGGAGCCTATGAAAGCCAGCAACCGCGCGGGG 1169
Db 6029 CGGAGCCTATGAAAGCCAGCAACCGCGCGGGG 5992

RESULT 6

US-08-899-575-170
; Sequence 170, Application US/08899575
; Patent No. 5804440
; GENERAL INFORMATION:
; APPLICANT: Burton, Dennis R
; APPLICANT: Barbas, Carlos F
; APPLICANT: Lerner, Richard A
; TITLE OF INVENTION: HUMAN NEUTRALIZING MONOCLONAL ANTIBODIES
; TITLE OF INVENTION: TO HUMAN IMMUNODEFICIENCY VIRUS
; NUMBER OF SEQUENCES: 170
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: The Scripps Research Institute, Office of
; ADDRESSEE: Patent Counsel
; STREET: 10666 No. 580440th Torrey Pines Road, Suite 220,
; STREET: Mail Drop TPC8
; CITY: La Jolla
; STATE: CA
; COUNTRY: USA
; ZIP: 92037
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/899,575
; FILING DATE: 24-JUL-1997
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/276,852
; FILING DATE: 18-JUL-1994
; APPLICATION NUMBER: US 08/178,302
; FILING DATE: 30-SEP-1993
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/954,148
; FILING DATE: 30-SEP-1992

ATTORNEY/AGENT INFORMATION:
NAME: Fitting, Thomas
REGISTRATION NUMBER: 34,163
REFERENCE/DOCKET NUMBER: SCRI452P
TELECOMMUNICATION INFORMATION:
TELEPHONE: 619-554-2937
TELEFAX: 619-554-6312
INFORMATION FOR SEQ ID NO: 170:
SEQUENCE CHARACTERISTICS:
LENGTH: 13254 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: circular
MOLECULE TYPE: DNA (genomic)
US-08-899-575-170

Query Match 31.9%; Score 454.8; DB 1; Length 13254;
Best Local Similarity 99.6%; Pred. No. 9.5e-133;
Matches 456; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
Qy 712 GTGTAGCGGTAGTTAGGCCACCACTTCAAGAACTCTGTAGCACCGCCTACATACCTCGCT 771
Db 6806 GTGTAGCGGTAGTTAGGCCACCACTTCAAGAACTCTGTAGCACCGCCTACATACCTCGCT 6865
Qy 772 CTGCTAATCTGTTTACCAAGTGGCTGCTGCCAGTGGCGGATAGTCTGTCTTTACCGGGTTG 831
Db 6866 CTGCTAATCTGTTTACCAAGTGGCTGCTGCCAGTGGCGGATAGTCTGTCTTTACCGGGTTG 6925
Qy 832 GACTCAAGACAGATAGTTACCGGATAAGCGGAGCGGCTGGGCTGAACGGGGGTTTCGTGC 891
Db 6926 GACTCAAGACAGATAGTTACCGGATAAGCGGAGCGGCTGGGCTGAACGGGGGTTTCGTGC 6985
Qy 892 ACACAGCCAGCTTGGAGGAGCACTACACCACTGAGTACCTACACGCTGAGCAT 951
Db 6986 ACACAGCCAGCTTGGAGGAGCACTACACCACTGAGTACCTACACGCTGAGCAT 7045
Qy 952 TGAGAAAGCGCCACGCTTCCCGAAGGAGGAGAAAGCGGACAGGTATCCGGTAAGCGGCAGG 1011
Db 7046 TGAGAAAGCGCCACGCTTCCCGAAGGAGGAGAAAGCGGACAGGTATCCGGTAAGCGGCAGG 7105
Qy 1012 GTCGGAACAGAGAGCGACAGAGGAGCTTCCAGGGGAAACGCTGATCTTTATAGT 1071
Db 7106 GTCGGAACAGAGAGCGACAGAGGAGCTTCCAGGGGAAACGCTGATCTTTATAGT 7165
Qy 1072 CCTCTCGGTTTCCGACCTCTGACTTGAGGCTGAGTCTGATCTCTCAGGGGG 1131
Db 7166 CCTCTCGGTTTCCGACCTCTGACTTGAGGCTGAGTCTGATCTCTCAGGGGG 7225
Qy 1132 CGGAGCCTATGGAAGAAAGCGGACCAAGCGGCGCGGGG 1169
Db 7226 CGGAGCCTATGGAAGAAAGCGGACCAAGCGGCGCGGG 7263

RESULT 7
PCT-US95-08743-156/c
Sequence 156, Application PC/TUS9508743
GENERAL INFORMATION:
APPLICANT:
TITLE OF INVENTION: HUMAN NEUTRALIZING MONOCLONAL ANTIBODIES
TO HUMAN IMMUNODEFICIENCY VIRUS
NUMBER OF SEQUENCES: 170
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25 (EPO)
CURRENT APPLICATION DATA:
APPLICATION NUMBER: PCT/US95/08743
FILING DATE: 11-JUL-1995
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/276,852
FILING DATE: 18-JUL-1994

INFORMATION FOR SEQ ID NO: 156:
SEQUENCE CHARACTERISTICS:
LENGTH: 13254 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: circular
MOLECULE TYPE: DNA (genomic)
PCT-US95-08743-156

Query Match 31.9%; Score 454.8; DB 5; Length 13254;
Best Local Similarity 99.6%; Pred. No. 9.5e-133;
Matches 456; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
Qy 712 GTGTAGCGGTAGTTAGGCCACCACTTCAAGAACTCTGTAGCACCGCCTACATACCTCGCT 771
Db 6449 GTGTAGCGGTAGTTAGGCCACCACTTCAAGAACTCTGTAGCACCGCCTACATACCTCGCT 6390
Qy 772 CTGCTAATCTGTTTACCAAGTGGCTGCTGCCAGTGGCGGATAAGTCTGTCTTTACCGGGTTG 831
Db 6389 CTGCTAATCTGTTTACCAAGTGGCTGCTGCCAGTGGCGGATAAGTCTGTCTTTACCGGGTTG 6330
Qy 832 GACTCAAGACAGATAGTTACCGGATAAGCGGAGCGGCTGGGCTGAACGGGGGTTTCGTGC 891
Db 6329 GACTCAAGACAGATAGTTACCGGATAAGCGGAGCGGCTGGGCTGAACGGGGGTTTCGTGC 6270
Qy 892 ACACAGCCAGCTTGGAGGAGCACTACACCACTGAGTACCTACACGCTGAGCAT 951
Db 6269 ACACAGCCAGCTTGGAGGAGCACTACACCACTGAGTACCTACACGCTGAGCAT 6210
Qy 952 TGAGAAAGCGCCACGCTTCCCGAAGGAGGAGAAAGCGGACAGGTATCCGGTAAGCGGCAGG 1011
Db 6209 TGAGAAAGCGCCACGCTTCCCGAAGGAGGAGAAAGCGGACAGGTATCCGGTAAGCGGCAGG 6150
Qy 1012 GTCGGAACAGAGAGCGACAGAGGAGCTTCCAGGGGAAACGCTGATCTTTATAGT 1071
Db 6149 GTCGGAACAGAGAGCGACAGAGGAGCTTCCAGGGGAAACGCTGATCTTTATAGT 6090
Qy 1072 CCTCTCGGTTTCCGACCTCTGACTTGAGGCTGAGTCTGATCTCTCAGGGGG 1131
Db 6089 CCTCTCGGTTTCCGACCTCTGACTTGAGGCTGAGTCTGATCTCTCAGGGGG 6030
Qy 1132 CGGAGCCTATGGAAGAAAGCGGACCAAGCGGCGCGGGG 1169
Db 6029 CGGAGCCTATGGAAGAAAGCGGACCAAGCGGCGCGGG 5992

RESULT 8
PCT-US95-08743-170
Sequence 170, Application PC/TUS9508743
GENERAL INFORMATION:
APPLICANT:
TITLE OF INVENTION: HUMAN NEUTRALIZING MONOCLONAL ANTIBODIES
TO HUMAN IMMUNODEFICIENCY VIRUS
NUMBER OF SEQUENCES: 170
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25 (EPO)
CURRENT APPLICATION DATA:
APPLICATION NUMBER: PCT/US95/08743
FILING DATE: 11-JUL-1995
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/276,852
FILING DATE: 18-JUL-1994
INFORMATION FOR SEQ ID NO: 170:
SEQUENCE CHARACTERISTICS:
LENGTH: 13254 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: circular
MOLECULE TYPE: DNA (genomic)

PCT-US95-08743-170

```

Query Match          31.9%; Score 454.8; DB 5; Length 13254;
Best Local Similarity 99.6%; Pred. No. 9.5e-133;
Matches 456; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 712 GTGTAGCCCTAGTTAGGCGCACCTTCAAGAACTCTGTAGCACCGCTTACATACCTCGCT 771
Db 6806 GTGTAGCCCTAGTTAGGCGCACCTTCAAGAACTCTGTAGCACCGCTTACATACCTCGCT 6865
QY 772 CTGCTAATCTCTTACCACTGCTCTGCCAGTGGCGATAAGTCTGTACCGGGTTG 831
Db 6866 CTGCTAATCTCTTACCACTGCTCTGCCAGTGGCGATAAGTCTGTACCGGGTTG 6925
QY 832 GACTCAAGACGATAGTTACCGGATAAGCGCGAGCGGCTCGGCTGAACGGGGTTCGTGC 891
Db 6926 GACTCAAGACGATAGTTACCGGATAAGCGCGAGCGGCTCGGCTGAACGGGGTTCGTGC 6985
QY 892 ACACAGCCCGAGCTTGGCGGAGCACTTACACCGAACTGAGATACCTACAGCGTGAGCAT 951
Db 6986 ACACAGCCCGAGCTTGGCGGAGCACTTACACCGAACTGAGATACCTACAGCGTGAGCAT 7045
QY 952 TCAGAAAGCGCCAGCTTCCCGAAGGAGAGAAAGCGGACAGGTATCCGGTAAGCGGCAGG 1011
Db 7046 TCAGAAAGCGCCAGCTTCCCGAAGGAGAGAAAGCGGACAGGTATCCGGTAAGCGGCAGG 7105
QY 1012 GTCGGAACAGAGAGCGGACAGGAGGAGCTTCCAGGGGAAACGCTCGGTATCTTTATAGT 1071
Db 7106 GTCGGAACAGAGAGCGGACAGGAGGAGCTTCCAGGGGAAACGCTCGGTATCTTTATAGT 7165
QY 1072 CCTGCGGGTTTCCCGACCTCTGACTTGAGCGTGCATTTTGTGATGCTCGTCAGGGGG 1131
Db 7166 CCTGCGGGTTTCCCGACCTCTGACTTGAGCGTGCATTTTGTGATGCTCGTCAGGGGG 7225
QY 1132 CGGAGCCTATGAAAAAGCGCAGCAACCGCGCGCGGGG 1169
Db 7226 CGGAGCCTATGAAAAAGCGCAGCAACCGCGCGCGGAG 7263

```

```

RESULT 9
US-08-594-469-9/c
; Sequence 9, Application US/08594469
; Patent No. 5700665
; GENERAL INFORMATION:
; APPLICANT: LEGOUX, Richard
; APPLICANT: MALDONADO, Paul
; APPLICANT: SALOME, Marc
; TITLE OF INVENTION: Method for the extraction of
; TITLE OF INVENTION: periplasmic proteins of prokaryotic microorganisms in the
; NUMBER OF SEQUENCES: 9
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Bacon & Thomas
; STREET: 625 Slaters Lane - Fourth Floor
; CITY: Alexandria
; STATE: Virginia
; COUNTRY: USA
; ZIP: 22314
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/594,469
; FILING DATE:
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: FR 95 01083
; FILING DATE: 31-JAN-1995
; ATTORNEY/AGENT INFORMATION:
; NAME: FICHTER, Richard E

```

```

; REGISTRATION NUMBER: 26,382
; REFERENCE/DOCKET NUMBER: REF/LEGOUX
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (703) 683-0500
; TELEFAX: (703) 683-1080
; INFORMATION FOR SEQ ID NO: 9:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1905 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; US-08-594-469-9

Query Match          31.8%; Score 453.8; DB 1; Length 1905;
Best Local Similarity 99.6%; Pred. No. 6.3e-133;
Matches 455; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 712 GTGTAGCCCTAGTTAGGCGCACCTTCAAGAACTCTGTAGCACCGCTTACATACCTCGCT 771
Db 518 GTGTAGCCCTAGTTAGGCGCACCTTCAAGAACTCTGTAGCACCGCTTACATACCTCGCT 459
QY 772 CTGCTAATCTCTTACCACTGCTCTGCCAGTGGCGATAAGTCTGTACCGGGTTG 831
Db 458 CTGCTAATCTCTTACCACTGCTCTGCCAGTGGCGATAAGTCTGTACCGGGTTG 399
QY 832 GACTCAAGACGATAGTTACCGGATAAGCGCGAGCGGCTCGGCTGAACGGGGTTCGTGC 891
Db 398 GACTCAAGACGATAGTTACCGGATAAGCGCGAGCGGCTCGGCTGAACGGGGTTCGTGC 339
QY 892 ACACAGCCCGAGCTTGGAGGAGCACTTACACCGAACTGAGATACCTACAGCGTGAGCAT 951
Db 338 ACACAGCCCGAGCTTGGAGGAGCACTTACACCGAACTGAGATACCTACAGCGTGAGCAT 279
QY 952 TGAGAAAGCGCCAGCTTCCCGAAGGAGAGAAAGCGGACAGGTATCCCGTAAGCGGCAGG 1011
Db 278 TGAGAAAGCGCCAGCTTCCCGAAGGAGAGAAAGCGGACAGGTATCCCGTAAGCGGCAGG 219
QY 1012 GTCGGAACAGAGAGCGGACGAGGAGGAGCTTCCAGGGGAAACGCTCGGTATCTTTATAGT 1071
Db 218 GTCGGAACAGAGAGCGGACGAGGAGGAGCTTCCAGGGGAAACGCTCGGTATCTTTATAGT 159
QY 1072 CCTGTCGGGTTTCCCGACCTCTGACTTGAGCGTGCATTTTGTGATGCTCGTCAGGGGG 1131
Db 158 CCTGTCGGGTTTCCCGACCTCTGACTTGAGCGTGCATTTTGTGATGCTCGTCAGGGGG 99
QY 1132 CGGAGCCTATGAAAAAGCGCAGCAACCGCGCGCGGG 1168
Db 98 CGGAGCCTATGAAAAAGCGCAGCAACCGCGCGCGGAG 62

```

```

RESULT 10
US-08-906-957-9/c
; Sequence 9, Application US/08906957
; Patent No. 5856142
; GENERAL INFORMATION:
; APPLICANT: LEGOUX, Richard
; APPLICANT: MALDONADO, Paul
; APPLICANT: SALOME, Marc
; TITLE OF INVENTION: Method for the extraction of
; TITLE OF INVENTION: periplasmic proteins of prokaryotic microorganisms in the
; NUMBER OF SEQUENCES: 9
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Bacon & Thomas
; STREET: 625 Slaters Lane - Fourth Floor
; CITY: Alexandria
; STATE: Virginia
; COUNTRY: USA
; ZIP: 22314
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk

```

```

:
: COMPUTER: IBM PC compatible
: OPERATING SYSTEM: PC-DOS/MS-DOS
: SOFTWARE: Patent Release #1.0, Version 1.0
: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: US/08/906,957
: FILING DATE: 06-AUG-1997
: CLASSIFICATION: 435
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER: 08/594,469
: FILING DATE:
: APPLICATION NUMBER: FR 95 01083
: FILING DATE: 31-JAN-1995
: ATTORNEY/AGENT INFORMATION:
: NAME: FICHTER, Richard E
: REGISTRATION NUMBER: 26,382
: REFERENCE/DOCKET NUMBER: REF/LEGOUX
: TELECOMMUNICATION INFORMATION:
: TELEPHONE: (703) 683-0500
: TELEFAX: (703) 683-1080
: INFORMATION FOR SEQ ID NO: 9:
: SEQUENCE CHARACTERISTICS:
: LENGTH: 1905 base pairs
: TYPE: nucleic acid
: STRANDEDNESS: single
: TOPOLOGY: linear
: MOLECULE TYPE: DNA (genomic)
:
: US-08-906-957-9

```

Query Match	31.88;	Score 453.8;	DB 2;	Length 1905;
Best Local Similarity	99.68;	Pred. No. 6.3e-133;		
Matches 455;	Conservative 0;	Mismatches 2;	Indels 0;	Gaps 0;
Qy	712	GTGTAGCCGTAGTTAGGCCACACACTTCAAGAACTCTGTAGCACCGCTACATACCTCGCT	771	
Db	518	GTGTAGCCGTAGTTAGGCCACCACTTCAAGAACTCTGTAGCACCGCTACATACCTCGCT	459	
Qy	772	CTGCTAATCTGTTTACCAGTGGCTGCCAGTGGCGATAAGTGTGTCTTTACCGGGTTG	831	
Db	458	CTGCTAATCTGTTTACCAGTGGCTGCCAGTGGCGATAAGTGTGTCTTTACCGGGTTG	399	
Qy	832	GACTCAAGACGATAGTTTACCGGATAAGCGCAGCGGTCCGGCTGAACGGGGGTTCTGTGC	891	
Db	398	GACTCAAGACGATAGTTTACCGGATAAGCGCAGCGGTCCGGCTGAACGGGGGTTCTGTGC	339	
Qy	892	ACACAGCCACGCTTGGAGCGAACCACCTACACCAACTGAGATACCTACACGCTGAGCAT	951	
Db	338	ACACAGCCACGCTTGGAGCGAACCACCTACACCAACTGAGATACCTACACGCTGAGCAT	279	
Qy	952	TGAGAAAGCGCCACGCTTCCGGAAGGGAGAAAGCGGACAGGTATCCGGTAAGCGGCAGG	1011	
Db	278	TGAGAAAGCGCCACGCTTCCGGAAGGGAGAAAGCGGACAGGTATCCGGTAAGCGGCAGG	219	
Qy	1012	GTGCGAAACAGGAGACGACGAGGAGGCTTCCAGGGGGAACGCGCTGTATCTTTATAGT	1071	
Db	218	GTGCGAAACAGGAGACGACGAGGAGGCTTCCAGGGGGAACGCGCTGTATCTTTATAGT	159	
Qy	1072	CTGTTCGGGTTTCGCCACCTCTGTACTTGTAGCGTCGATTTTGTGTATGCTCTCAGGGGGG	1131	
Db	158	CCTGTCCGGTTTCGCCACCTCTGTACTTGTAGCGTCGATTTTGTGTATGCTCTCAGGGGGG	99	
Qy	1132	CGAGCCCTATGGAAAAACGCCAGCAACGGCCCGGGG	1168	
Db	98	CGAGCCCTATGGAAAAACGCCAGCAACGGCCCGCGAG	62	

RESULT 11
US-08-594-469-1/c
; Sequence 1, Application US/08594469
; Patent No. 5700665
; GENERAL INFORMATION:
; APPLICANT: LEGOUX, Richard
; APPLICANT: MALDONADO, Paul

APPLICANT: SALOME, Marc
 TITLE OF INVENTION: Method for the extraction of
 TITLE OF INVENTION: periplasmic proteins of prokaryotic microorganisms in the
 TITLE OF INVENTION: presence of arginine
 NUMBER OF SEQUENCES: 9
 CORRESPONDENCE ADDRESS:
 ADDRESSEE: Bacon & Thomas
 STREET: 625 Slaters Lane - Fourth Floor
 CITY: Alexandria
 STATE: Virginia
 COUNTRY: USA
 ZIP: 22314
 COMPUTER READABLE FORM:
 MEDIUM TYPE: Floppy disk
 COMPUTER: IBM PC compatible
 OPERATING SYSTEM: PC-DOS/MS-DOS
 SOFTWARE: PatentIn Release #1.0, Version #1.25
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/08/594,469
 FILING DATE:
 CLASSIFICATION: 435
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: FR 95 01083
 FILING DATE: 31-JAN-1995
 ATTORNEY/AGENT INFORMATION:
 NAME: FICHTER, Richard E
 REGISTRATION NUMBER: 26,382
 REFERENCE/DOCKET NUMBER: REF/LEGUOX
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: (703) 683-0500
 TELEFAX: (703) 683-1080
 INFORMATION FOR SEQ ID NO: 1:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 4410 base pairs
 TYPE: nucleic acid
 STRANDEDNESS: single
 TOPOLOGY: linear
 MOLECULE TYPE: DNA (genomic)
 US-08-594-469-1

Query Match	31.8%;	Score 453.8;	DB 1;	Length 4410;
Best Local Similarity	99.6%;	Pred. No. 1e-132;		
Matches 455;	Conservative	2;	Indels	0; Gaps
Qy	712	GTGTAGCCGTTAGTTAGGCCACACCTTCAAGAACTCTGTAGCACCGCCCTACATACCTCGCT	771	
Db	3023	GTGTAGCCGTTAGTTAGGCCACCACCTTCAAGAACTCTGTAGCACCGCCCTACATACCTCGCT	2964	
Qy	772	CTGCTAAATCTGTTTACCAGTGGCTGCTGCCAGTGGCGATAAGTGTGTCTTACCGGGTTG	831	
Db	2963	CTGCTAAATCTGTTTACCAGTGGCTGCTGCCAGTGGCGATAAGTGTGTCTTACCGGGTTG	2904	
Qy	832	GACTCAACACCATAGTTTACCGSATAAAGCGCAGCGCTCGGGCTGAACGGGGGTTCTCGTC	891	
Db	2903	GACTCAACACCATAGTTTACCGSATAAAGCGCAGCGCTCGGGCTGAACGGGGGTTCTCGTC	2844	
Qy	892	ACACAGCCCGCTTGGAGCGAAGCACTTACACCGAACTTGAGATACCTACACGCGTGAGCAT	951	
Db	2843	ACACAGCCCGCTTGGAGCGAAGCACTTACACCGAACTTGAGATACCTACACGCGTGAGCAT	2784	
Qy	952	TGAGAAACGCCACGCTTCCGGAAGGGAGAAAGCGCGACAGGTATCCCGGTAAAGCGGCAGG	1011	
Db	2783	TGAGAAACGCCACGCTTCCGGAAGGGAGAAAGCGCGACAGGTATCCCGGTAAAGCGGCAGG	2724	
Qy	1012	GTGCGAAACAGGAGACGCGACGAGGAGCTTCCAGGGGGAACCGCTGTATCTTTATAGT	1071	
Db	2723	GTGCGAAACAGGAGACGCGACGAGGAGCTTCCAGGGGGAACCGCTGTATCTTTATAGT	2664	
Qy	1072	CCTGTGCGGGTTTGGCCACCTCTGACTTGTAGCGTTCGATTTTTTGTGATGCTGTCAGGGGGG	1131	
Db	2663	CCTGTGCGGGTTTGGCCACCTCTGACTTGTAGCGTTCGATTTTTTGTGATGCTGTCAGGGGGG	2604	

QY 1132 CGAGGCTATGGAACACCCAGCAACGCGCGGGG 1168
 ||||||||||||||||||||||||||||||||||||
 Db 2603 CGAGGCTATGGAACACCCAGCAACGCGCGGAG 2567

RESULT 12

US-08-906-957-1/c
 ; Sequence 1, Application US/08906957
 ; Patent No. 5856142
 ; GENERAL INFORMATION:
 ; APPLICANT: LEGOUX, Richard
 ; APPLICANT: MALDONADO, Paul
 ; APPLICANT: SALONE, Marc
 ; TITLE OF INVENTION: Method for the extraction of
 ; TITLE OF INVENTION: periplasmic proteins of prokaryotic microorganisms in the
 ; TITLE OF INVENTION: presence of arginine
 ; NUMBER OF SEQUENCES: 9
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: Bacon & Thomas
 ; STREET: 625 Slaters Lane - Fourth Floor
 ; CITY: Alexandria
 ; STATE: Virginia
 ; COUNTRY: USA
 ; ZIP: 22314
 ; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: Floppy disk
 ; COMPUTER: IBM PC compatible
 ; OPERATING SYSTEM: PC-DOS/MS-DOS
 ; SOFTWARE: PatentIn Release #1.0, Version #1.25
 ; CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: US/08/906,957
 ; FILING DATE: 06-AUG-1997
 ; CLASSIFICATION: 435
 ; PRIOR APPLICATION DATA:
 ; APPLICATION NUMBER: 08/594,469
 ; FILING DATE:
 ; APPLICATION NUMBER: FR 95 01083
 ; FILING DATE: 31-JAN-1995
 ; ATTORNEY/AGENT INFORMATION:
 ; NAME: FICHTER, Richard E
 ; REGISTRATION NUMBER: 26,382
 ; REFERENCE/DOCKET NUMBER: REF/LEGOUX
 ; TELECOMMUNICATION INFORMATION:
 ; TELEPHONE: (703) 683-0500
 ; TELEFAX: (703) 683-1080
 ; INFORMATION FOR SEQ ID NO: 1:
 ; SEQUENCE CHARACTERISTICS:
 ; LENGTH: 4410 base pairs
 ; TYPE: nucleic acid
 ; STRANDEDNESS: single
 ; TOPOLOGY: linear
 ; MOLECULE TYPE: DNA (genomic)
 ; US-08-906-957-1

Query Match 31.8%; Score 453.8; DB 2; Length 4410;
 Best Local Similarity 99.6%; Pred. No. 1e-132;
 Matches 455; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
 QY 712 GTGTAGCCGTAGTACGACCACTTCAAGAACTGTAGCACCGCCTACATACCTCGCT 771
 ||||||||||||||||||||||||||||||||||||
 Db 3023 GTGTAGCCGTAGTACGACCACTTCAAGAACTGTAGCACCGCCTACATACCTCGCT 2964
 ||||||||||||||||||||||||||||||||||||
 QY 772 CTGCTAATCTGTTACAGTGGCTGCTGCAGTGGGATAAGTCTGTCTTACCGGGTTG 831
 ||||||||||||||||||||||||||||||||||||
 Db 2963 CTGCTAATCTGTTACAGTGGCTGCTGCAGTGGGATAAGTCTGTCTTACCGGGTTG 2904
 ||||||||||||||||||||||||||||||||||||
 QY 832 GACTCAAGCAGTAGTACCGGATAAGCGCAGCGGTGGGTGAACGGGGGTTTCGTGC 891
 ||||||||||||||||||||||||||||||||||||
 Db 2903 GACTCAAGCAGTAGTACCGGATAAGCGCAGCGGTGGGTGAACGGGGGTTTCGTGC 2844
 ||||||||||||||||||||||||||||||||||||
 QY 892 ACACAGCCCGAGCTTGAGCGGACGACCTACACGAACTGAGATACCTACAGCGTGAAGAT 951
 ||||||||||||||||||||||||||||||||||||

Db 2843 ACACAGCCCGAGCTTGGAGCGGAACCACTACACCAACTGAGATACCTACAGCGTGAGCAT 2784
 ||||||||||||||||||||||||||||||||||||
 QY 952 TGAGAAAGCGCCAGCGCTTCCCGAAGGGAGAAAGCGGACAGGTATCCGGTAAGCGCGAGG 1011
 ||||||||||||||||||||||||||||||||||||
 Db 2783 TGAGAAAGCGCCAGCGCTTCCCGAAGGGAGAAAGCGGACAGGTATCCGGTAAGCGCGAGG 2724
 ||||||||||||||||||||||||||||||||||||
 QY 1012 GTCGGAACAGGAGCGGACAGGAGGAGCTTCCAGGGGAAACGCGCTGTATCTTTATAGT 1071
 ||||||||||||||||||||||||||||||||||||
 Db 2723 GTCGGAACAGGAGCGGACAGGAGGAGCTTCCAGGGGAAACGCGCTGTATCTTTATAGT 2664
 ||||||||||||||||||||||||||||||||||||
 QY 1072 CCTGTGCGGGTTCCGCCACCTCTGACTTGGAGCTGATTTTGTGATGCTGCTCAGGGGGG 1131
 ||||||||||||||||||||||||||||||||||||
 Db 2663 CCTGTGCGGGTTCCGCCACCTCTGACTTGGAGCTGATTTTGTGATGCTGCTCAGGGGGG 2604
 ||||||||||||||||||||||||||||||||||||
 QY 1132 CGGAGCCTATGGAACCAACGCGGAGCAACGCGCGGGG 1168
 ||||||||||||||||||||||||||||||||||||
 Db 2603 CGGAGCCTATGGAACCAACGCGGAGCAACGCGCGGGG 2567

RESULT 13

PCT-US91-02954-12/c
 ; Sequence 12, Application PC/TUS9102954
 ; GENERAL INFORMATION:
 ; APPLICANT: PEPINSKY, R. BLAKE
 ; APPLICANT: ROSA, MARGARET D.
 ; APPLICANT: STOSSEL, THOMAS P.
 ; TITLE OF INVENTION: MULTIMERIC GELSOLIN FUSION CONSTRUCTS
 ; NUMBER OF SEQUENCES: 14
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: FISH & NEAVE
 ; STREET: 875 Third Avenue
 ; CITY: New York
 ; STATE: New York
 ; COUNTRY: United States of America
 ; ZIP: 10022
 ; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: Floppy disk
 ; COMPUTER: IBM PC compatible
 ; OPERATING SYSTEM: PC-DOS/MS-DOS
 ; SOFTWARE: PatentIn Release #1.0, Version #1.25
 ; CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: PCT/US91/02954
 ; FILING DATE: 19910503
 ; CLASSIFICATION: 436
 ; PRIOR APPLICATION DATA:
 ; APPLICATION NUMBER: US 07/520,368
 ; FILING DATE: 04-MAY-1990
 ; ATTORNEY/AGENT INFORMATION:
 ; NAME: Haley Jr., James F.
 ; REGISTRATION NUMBER: 27,794
 ; REFERENCE/DOCKET NUMBER: B144CIP
 ; TELECOMMUNICATION INFORMATION:
 ; TELEPHONE: (212) 715-0600
 ; TELEFAX: (212) 715-0634
 ; TELEX: 14-8367
 ; INFORMATION FOR SEQ ID NO: 12:
 ; SEQUENCE CHARACTERISTICS:
 ; LENGTH: 6151 base pairs
 ; TYPE: NUCLEIC ACID
 ; STRANDEDNESS: double
 ; TOPOLOGY: circular
 ; PCT-US91-02954-12

Query Match 31.8%; Score 453.8; DB 5; Length 6151;
 Best Local Similarity 99.6%; Pred. No. 1.e-132;
 Matches 455; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
 QY 712 GTGTAGCCGTAGTACGACCACTTCAAGAACTGTAGCACCGCCTACATACCTCGCT 771
 ||||||||||||||||||||||||||||||||||||
 Db 4761 GTGTAGCCGTAGTACGACCACTTCAAGAACTGTAGCACCGCCTACATACCTCGCT 4702
 ||||||||||||||||||||||||||||||||||||
 QY 772 CTGCTAATCTGTTACAGTGGCTGCTGCCAGTGGGATAGTCTGTCTTACCGGGTTG 831

|||||
Db 4701 CTGCTAATCTGTACAGTGGCTGCTCCAGTGGCGGATAAGTCTGTCTTACCGGGTTG 4642
Qy 832 GACTCAAGACGATAGTTACCGGATAAGCGCAGCGGTCGGGCTGAACCGGGGGTTCTGTCG 891
Db 4641 GACTCAAGACGATAGTTACCGGATAAGCGCAGCGGTCGGGCTGAACCGGGGGTTCTGTCG 4582
Qy 892 ACACAGCCAGCTTGGAGCGAACGACCTTACACCGAACTGAGATACCTACACGCTGAGCAT 951
Db 4581 ACACAGCCAGCTTGGAGCGAACGACCTTACACCGAACTGAGATACCTACACGCTGAGCAT 4522
Qy 952 TGAGAAAGCCGACGCTTCCCGAAGGAGGAGAAAGCGGACAGGTATCCGGTAAGCGGCAGG 1011
Db 4521 TGAGAAAGCCGACGCTTCCCGAAGGAGGAGAAAGCGGACAGGTATCCGGTAAGCGGCAGG 4462
Qy 1012 GTCGGAACAGGAGCGCACGAGGAGCTTCCAGGGGAAACGCTGTGTATCTTTATAGT 1071
Db 4461 GTCGGAACAGGAGCGCACGAGGAGCTTCCAGGGGAAACGCTGTGTATCTTTATAGT 4402
Qy 1072 CCTGTCCGGTTTCGCCACCTCTGACTTGAGCGTCGATTTTGTGATGCTCGTCAGGGGGG 1131
Db 4401 CCTGTCCGGTTTCGCCACCTCTGACTTGAGCGTCGATTTTGTGATGCTCGTCAGGGGGG 4342
Qy 1132 CGGAGCCTATGGAAGAACCGCAGCAACCGCGCCGGG 1168
Db 4341 CGGAGCCTATGGAAGAACCGCAGCAACCGCGCCGGG 4305

RESULT 14
US-08-460-343B-1/c
; Sequence 1, Application US/08460343B
; Patent No. 5741664
; GENERAL INFORMATION:
; APPLICANT: Marcus D. Ballinger and James A. Wells
; TITLE OF INVENTION: SUBSTITUTED VARIANTS CAPABLE OF CLEAVING
; TITLE OF INVENTION: SUBSTRATES CONTAINING DIBASIC RESIDUES
; NUMBER OF SEQUENCES: 74
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Genentech, Inc.
; STREET: 1 DNA Way
; CITY: South San Francisco
; STATE: California
; COUNTRY: USA
; ZIP: 94080
; COMPUTER READABLE FORM:
; MEDIUM TYPE: 3.5 inch, 1.44 Mb floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: WinPatIn (Genentech)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/460,343B
; FILING DATE: 01-Jun-1995
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/398028
; FILING DATE: 03-mar-1995
; ATTORNEY/AGENT INFORMATION:
; NAME: Kubinec, Jeffrey S.
; REGISTRATION NUMBER: 36,575
; REFERENCE/DOCKET NUMBER: P0936C1
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 650/225-8228
; TELEFAX: 650/952-9881
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 8119 base pairs
; TYPE: Nucleic Acid
; STRANDEDNESS: Single
; TOPOLOGY: Linear
US-08-460-343B-1

Query Match 31.8%; Score 453.8; DB 1; Length 8119;

Best Local Similarity 99.6%; Pred. No. 1.5e-132;
Matches 455; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
Qy 712 GTGTAGCCGTAGTTAGGCGACACCTTCAAGAACTCTGTAGCACCCCTACATACCTCGCT 771
Db 2882 GTGTAGCCGTAGTTAGGCGACACCTTCAAGAACTCTGTAGCACCCCTACATACCTCGCT 2823
Qy 772 CTGCTAATCTGTACAGTGGCTGCTCCAGTGGCGGATAAGTCTGTCTTACCGGGTTG 831
Db 2822 CTGCTAATCTGTACAGTGGCTGCTCCAGTGGCGGATAAGTCTGTCTTACCGGGTTG 2763
Qy 832 GACTCAAGACGATAGTTACCGGATAAGCGCAGCGGTCGGGCTGAACCGGGGGTTCTGTCG 891
Db 2762 GACTCAAGACGATAGTTACCGGATAAGCGCAGCGGTCGGGCTGAACCGGGGGTTCTGTCG 2703
Qy 892 ACACAGCCAGCTTGGAGCGAACGACCTTACACCGAACTGAGATACCTACACGCTGAGCAT 951
Db 2702 ACACAGCCAGCTTGGAGCGAACGACCTTACACCGAACTGAGATACCTACACGCTGAGCAT 2643
Qy 952 TGAGAAAGCCGACGCTTCCCGAAGGAGGAGAAAGCGGACAGGTATCCGGTAAGCGGCAGG 1011
Db 2642 TGAGAAAGCCGACGCTTCCCGAAGGAGGAGAAAGCGGACAGGTATCCGGTAAGCGGCAGG 2583
Qy 1012 GTCGGAACAGGAGCGCACGAGGAGCTTCCAGGGGAAACGCTGTGTATCTTTATAGT 1071
Db 2582 GTCGGAACAGGAGCGCACGAGGAGCTTCCAGGGGAAACGCTGTGTATCTTTATAGT 2523
Qy 1072 CCTGTCCGGTTTCGCCACCTCTGACTTGAGCGTCGATTTTGTGATGCTCGTCAGGGGGG 1131
Db 2522 CCTGTCCGGTTTCGCCACCTCTGACTTGAGCGTCGATTTTGTGATGCTCGTCAGGGGGG 2463
Qy 1132 CGGAGCCTATGGAAGAACCGCAGCAACCGCGCCGGG 1168
Db 2462 CGGAGCCTATGGAAGAACCGCAGCAACCGCGCCGGG 2426

RESULT 15
US-08-398-028B-1/c
; Sequence 1, Application US/08398028B
; Patent No. 5780285
; GENERAL INFORMATION:
; APPLICANT: Marcus D. Ballinger and James A. Wells
; TITLE OF INVENTION: SUBSTITUTED VARIANTS CAPABLE OF CLEAVING
; TITLE OF INVENTION: SUBSTRATES CONTAINING DIBASIC RESIDUES
; NUMBER OF SEQUENCES: 74
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Genentech, Inc.
; STREET: 1 DNA Way
; CITY: South San Francisco
; STATE: California
; COUNTRY: USA
; ZIP: 94080
; COMPUTER READABLE FORM:
; MEDIUM TYPE: 3.5 inch, 1.44 Mb floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: WinPatIn (Genentech)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/398,028B
; FILING DATE: 03-Mar-1995
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Kubinec, Jeffrey S.
; REGISTRATION NUMBER: 36,575
; REFERENCE/DOCKET NUMBER: P0936
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 650/225-8228
; TELEFAX: 650/952-9881
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 8119 base pairs
; TYPE: Nucleic Acid
; STRANDEDNESS: Single

TOPOLOGY: Linear
US-08-398-028B-1

```

Query Match          31.8%; Score 453.8; DB 1; Length 8119;
Best Local Similarity 99.6%; Pred. No. 1.5e-132;
Matches 455; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 712 GTGTAGCCGTAGTTAGGACCACCTTCAAGAACTCTGTAGCACCGCTACATACCTCGCT 771
Db 2882 GTGTAGCCGTAGTTAGGACCACCTTCAAGAACTCTGTAGCACCGCTACATACCTCGCT 771
QY 772 CTGCTAATCCTGTTACCACTGCTCTGCCAGTGGCGATAGTCTGTCTTACCGGGTTG 831
Db 2822 CTGCTAATCCTGTTACCACTGCTCTGCCAGTGGCGATAGTCTGTCTTACCGGGTTG 831
QY 832 GACTCAAGACGATAGTTACCGGATAAGCGCAGCGCTCGGGCTCAACGGGGGTTGTCG 891
Db 2762 GACTCAAGACGATAGTTACCGGATAAGCGCAGCGCTCGGGCTCAACGGGGGTTGTCG 891
QY 892 ACACAGCCCGAGCTTGGAGCGAAGCACTACACCGAACTGAGATACCTACAGCGTGAGCAT 951
Db 2702 ACACAGCCCGAGCTTGGAGCGAAGCACTACACCGAACTGAGATACCTACAGCGTGAGCAT 951
QY 952 TGAGAAAGCGCCAGCGTTCCCGAAGGAGAAAGCGGACAGGTATCCGGTAAAGCGGCAGG 1011
Db 2642 TGAGAAAGCGCCAGCGTTCCCGAAGGAGAAAGCGGACAGGTATCCGGTAAAGCGGCAGG 1011
QY 1012 GTCGGAACAGGAGAGCCACGAGGAGCTTCCAGGGGGAAACCCCTGGTATCTTTATAGT 1071
Db 2582 GTCGGAACAGGAGAGCCACGAGGAGCTTCCAGGGGGAAACCCCTGGTATCTTTATAGT 1071
QY 1072 CCTGTGGGTTTCGCCACCTCTGACTTGAGCGTCTGATTTTCTGATGCTGTCAGGGGGG 1131
Db 2522 CCTGTGGGTTTCGCCACCTCTGACTTGAGCGTCTGATTTTCTGATGCTGTCAGGGGGG 1131
QY 1132 CGGAGCCTATGGAAACCGCCAGCAACGGCGCCGGG 1168
Db 2462 CGGAGCCTATGGAAACCGCCAGCAACGGCGCCGGAG 2426

```

Search completed: January 17, 2002, 11:53:09
Job time: 17316 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM nucleic - nucleic search, using sw model.

Run on: January 17, 2002, 08:54:03 ; Search time 6501.33 seconds
(without alignments)
2355.323 Million cell updates/sec

Title: US-09-242-202a-16
Perfect score: 1425
Sequence: 1 TGCCATGGCGGATTCCTT.....CAGCCTCTCCACAGGTACC 1425

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 11351937 seqs, 537289281 residues

Total number of hits satisfying chosen parameters: 22703874

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

EST:*
1: em_estfun:*
2: em_esthum:*
3: em_estin:*
4: em_estom:*
5: em_estpl:*
6: em_estba:*
7: em_estro:*
8: em_estov:*
9: em_hic:*
10: gb_estl:*
11: gb_est2:*
12: gb_hic:*
13: gb_gss:*
14: em_gss_fun:*
15: em_gss_hum:*
16: em_gss_inv:*
17: em_gss_pln:*
18: em_gss_pro:*
19: em_gss_rod:*
20: em_gss_vrt:*
21: em_gss_other:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
C 1	450.2	31.6	615	10 AV735664	AV735664
C 2	449.8	31.6	527	10 AL043585	AL043585 DKF2p434G
C 3	449.8	31.6	571	10 AL044178	AL044178 DKF2p434P
C 4	449.8	31.6	579	10 AL043613	AL043613 DKF2p434H
C 5	449.8	31.6	616	10 AV735756	AV735756
C 6	449.8	31.6	617	10 AJ281661	AJ281661 4A3A-P8G1
C 7	449.8	31.6	628	11 BF381364	BF381364 ASIR0004
C 8	449.8	31.6	954	10 AL044364	AL044364 DKF2p434C
C 9	448.8	31.5	629	10 AL593919	AL593919
C 10	448.8	31.5	1004	10 AJ281480	AJ281480 4A3A-P4G8
C 11	448.2	31.5	703	10 AJ281437	AJ281437
C 12	448.2	31.5	800	10 AJ281449	AJ281449 4A3A-P4D5

C 13	447.8	31.4	568	10 AJ281376	AJ281376 4A3A-P2G2
C 14	447.8	31.4	579	10 AJ281320	AJ281320 4A3A-P1H1
C 15	447.8	31.4	700	10 AJ281616	AJ281616 4A3A-P8A1
C 16	446.6	31.3	1067	10 AU081137	AU081137 AU081137
C 17	445	31.2	808	10 AU176264	AU176264 AU176264
C 18	443.8	31.1	548	10 AJ281654	AJ281654 4A3A-P8F1
C 19	442.8	31.1	1070	10 AJ281552	AJ281552 4A3A-P6F1
C 20	441.8	31.0	526	10 AL043840	AL043840 DKF2p434A
C 21	441.2	31.0	498	10 AL039576	AL039576 DKF2p434D
C 22	438.8	30.8	615	10 AL044413	AL044413 DKF2p434E
C 23	438.4	30.8	741	11 BF299419	BF299419 24A-6-11
C 24	437.8	30.7	734	10 AL039459	AL039459 DKF2p434O
C 25	437.2	30.7	1089	10 AU081124	AU081124 AU081124
C 26	435.8	30.6	480	13 C111G3	AJ226213 Clona int
C 27	435.8	30.6	636	10 AJ281699	AJ281699 4A3A-P9E3
C 28	435.2	30.5	756	10 BE749097	BE749097 601123138
C 29	434.8	30.5	718	13 AG010489	AG010489 Homo sapi
C 30	433.6	30.4	959	10 BE749147	BE749147 601123315
C 31	431	30.2	759	10 BE749118	BE749118 601123194
C 32	429.6	30.1	840	10 BE749178	BE749178 601123444
C 33	429.2	30.1	757	11 BF863156	BF863156 963041F03
C 34	425.8	29.9	630	10 AL042334	AL042334 DKF2p434O
C 35	425.2	29.8	784	13 AQ876119	AQ876119 V133E1 mt
C 36	423.8	29.7	590	10 AV613078	AV613078 AV613078
C 37	422.6	29.7	795	13 AQ876011	AQ876011 V132B5 mt
C 38	418.6	29.4	690	13 AG009464	AG009464 Homo sapi
C 39	417.4	29.3	579	10 AV604761	AV604761 AV604761
C 40	417	29.3	540	10 AV604844	AV604844 AV604844
C 41	415.8	29.2	485	10 AL044354	AL044354 DKF2p434C
C 42	411	28.8	436	10 AL040947	AL040947 DKF2p434K
C 43	407.6	28.6	629	13 AG011367	AG011367 Homo sapi
C 44	403	28.3	794	13 AQ875988	AQ875988 V131H3 mt
C 45	401.6	28.2	706	10 BE268532	BE268532 601125429

ALIGNMENTS

RESULT 1
AV735664/c
LOCUS AV735664 615 bp mRNA EST 17-OCT-2000
DEFINITION AV735664 CB Homo sapiens CDNA clone CBNAMES07 5', mRNA sequence.
ACCESSION AV735664
VERSION AV735664.1 GI:10853245
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 615)
AUTHORS Zhang,Q., Ye,M., Wu,X., Gu,J., Huang,Q., Zhou,J., Shen,Y., Han,Z., Chen,S., Mao,M. and Chen,Z.
TITLE Homo sapiens CB library CDNA clones
JOURNAL Unpublished (2000)
COMMENT Contact: Zhu Chen
Shanghai Institute of Hematology, Rui-Jin Hospital
197 Rui-Jin II Road, Shanghai 200025, P. R. China
Tel: 86-21-64740490
Fax: 86-21-64743206
Email: mbshiem.stn.sh.cn
This clone is available at Shanghai Hematology Institute in Shanghai.
Chinese National Human Genome Center at Shanghai
351 Guo Shoujing Road, Zhangjiang Hi-Tech Park, Pudong.

FEATURES

source
1. 615
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="CBNAME07"
/clone_lib="CB"
/tissue_type="cord blood"
/cell_type="CD34+ hematopoietic stem/progenitor cell"
/lab_host="BM25.8"

BASE COUNT 129 a 176 c 159 g 147 t 4 Others

ORIGIN

/note="Vector: pBluescript; Site_1: EcoRI; The insert is
Cloned randomly with the EcoRI digestion"

Query Match 31.6%; Score 450.2; DB 10; Length 615;
Best Local Similarity 99.1%; Pred. No. 2.7e-116;
Matches 452; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 712 GTGTAGCCGTAGTTAGGCGACCACTTCAAGAACTCTGTAGCAGCGCTACATACCTCGCT 771
Db 460 GTGTAGCCGTAGTTAGGCGACCACTTCAAGAACTCTGTAGCAGCGCTACATACCTCGCT 771
QY 772 CTGCTAAATCCTGTTACCACTGCTGCTGCCAGTGGCGGATAAGTCTGTCTTACCGGTTG 831
Db 400 CTGCTAAATCCTGTTACCACTGCTGCTGCCAGTGGCGGATAAGTCTGTCTTACCGGTTG 831
QY 832 GACTCAAGACGATAGTTACCGGATAAGCGCAGCGTTCGGCTGGAACGGGGTTCGTTG 341
Db 340 GACTCAAGACGATAGTTACCGGATAAGCGCAGCGTTCGGCTGGAACGGGGTTCGTTG 341
QY 892 ACACAGCCAGCTTGGAGCGAAGCACTACACCGAACTGAGATACCTACAGCGTGCAGCAT 951
Db 280 ACACAGCCAGCTTGGAGCGAAGCACTACACCGAACTGAGATACCTACAGCGTGCAGCAT 951
QY 952 TGAGAAAGCCCGCTTCCGGAAGGGAAGAGCGGAGCTTCCAGGGGAAACGCTGTATCTTTATAGT 101
Db 220 TGAGAAAGCCCGCTTCCGGAAGGGAAGAGCGGAGCTTCCAGGGGAAACGCTGTATCTTTATAGT 101
QY 1012 GTCGGAACAGGAGCGCAGGAGGAGCTTCCAGGGGAAACGCTGTATCTTTATAGT 1071
Db 160 GTCGGAACAGGAGCGCAGGAGGAGCTTCCAGGGGAAACGCTGTATCTTTATAGT 1071
QY 1072 CTTGTCGGTTTCGGCCACCTCTGACTTGCAGCTCGATTTTGTGATCTCTGTCAGGGGG 1131
Db 100 CTTGTCGGTTTCGGCCACCTCTGACTTGCAGCTCGATTTTGTGATCTCTGTCAGGGGG 41
QY 1132 CGAGCGCTATGGAANAACGCCAGCAACGCCGCGG 1167
Db 40 CGAGCGCTATGGAANAACGCCAGCAACGCCGCGG 5

RESULT 2
AL043585/c
LOCUS
DEFINITION DKFZP434G0127_s1.434 (synonym: htes3) Homo sapiens cDNA clone
ACCESSION DKFZP434G0127.3, mRNA sequence.
VERSION AL043585
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 527)
AUTHORS Bloecker, H., Boecker, M., Brandt, P., Mewes, H.W., Gassenhuber, J. and Wiemann, S.
TITLE EST (Bloecker, et al.)
JOURNAL Unpublished (1999)
COMMENT On Jul 8, 1999 this sequence version replaced gi:5866785.
Contact: Bloecker H
MIPS

An Klopferspitz 18a D-82152 Martinsried, Germany
This is the 3' sequence of the clone insert
Clone from S. Wiemann, Molecular Genome Analysis, German Cancer
Research Center (DKFZ); Email s.wiemann@dkfz-heidelberg.de;
Braunschweig/Germany) within the cDNA sequencing consortium of the
German Genome Project.
No r1 sequence available.
This clone (DKFZP434G0127) is available at the RZPD in Berlin.
Please contact the RZPD: Ressourcenzentrum, Heubnerweg 6, 14059

FEATURES
Source
Location/Qualifiers
1..527

/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="DKFZP434G0127"
/clone_lib="434 (synonym: htes3)"
/tissue_type="testis"
/dev_stage="adult"
/lab_host="DH10B"

BASE COUNT 113 a 158 c 137 g 119 t

ORIGIN

/note="Vector: pSport1; Site_1: NotI; Site_2: SalI"

Query Match 31.6%; Score 449.8; DB 10; Length 527;
Best Local Similarity 99.6%; Pred. No. 3.3e-116;
Matches 451; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 712 GTGTAGCCGTAGTTAGGCGACCACTTCAAGAACTCTGTAGCAGCGCTACATACCTCGCT 771
Db 474 GTGTAGCCGTAGTTAGGCGACCACTTCAAGAACTCTGTAGCAGCGCTACATACCTCGCT 771
QY 772 CTGCTAAATCCTGTTACCACTGCTGCTGCCAGTGGCGGATAAGTCTGTCTTACCGGTTG 831
Db 414 CTGCTAAATCCTGTTACCACTGCTGCTGCCAGTGGCGGATAAGTCTGTCTTACCGGTTG 831
QY 832 GACTCAAGACGATAGTTACCGGATAAGCGCAGCGTTCGGCTGGAACGGGGTTCGTTG 891
Db 354 GACTCAAGACGATAGTTACCGGATAAGCGCAGCGTTCGGCTGGAACGGGGTTCGTTG 891
QY 892 ACACAGCCCGCTTGGAGCGAAGCACTACACCGAACTGAGATACCTACAGCGTGCAGCAT 951
Db 294 ACACAGCCCGCTTGGAGCGAAGCACTACACCGAACTGAGATACCTACAGCGTGCAGCAT 951
QY 952 TGAGAAAGCCCGCTTCCGGAAGGGAAGAGCGGAGCTTCCAGGGGAAACGCTGTATCTTTATAGT 1071
Db 234 TGAGAAAGCCCGCTTCCGGAAGGGAAGAGCGGAGCTTCCAGGGGAAACGCTGTATCTTTATAGT 1071
QY 1012 GTCGGAACAGGAGCGCAGGAGGAGCTTCCAGGGGAAACGCTGTATCTTTATAGT 1071
Db 174 GTCGGAACAGGAGCGCAGGAGGAGCTTCCAGGGGAAACGCTGTATCTTTATAGT 1071
QY 1072 CTTGTCGGTTTCGGCCACCTCTGACTTGCAGCTCGATTTTGTGATCTCTGTCAGGGGG 1131
Db 114 CTTGTCGGTTTCGGCCACCTCTGACTTGCAGCTCGATTTTGTGATCTCTGTCAGGGGG 55
QY 1132 CGAGCGCTATGGAANAACGCCAGCAACGCCGCGG 1164
Db 54 CGAGCGCTATGGAANAACGCCAGCAACGCCGCGG 22

RESULT 3
AL044178/c
LOCUS
DEFINITION DKFZP434P0828_s1.434 (synonym: htes3) Homo sapiens cDNA clone
ACCESSION DKFZP434P0828.3, mRNA sequence.
VERSION AL044178
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 571)
AUTHORS Bloecker, H., Boecker, M., Brandt, P., Mewes, H.W., Gassenhuber, J. and Wiemann, S.
TITLE EST (Bloecker, et al.)
JOURNAL Unpublished (1999)
COMMENT On Jul 9, 1999 this sequence version replaced gi:5866789.
Contact: Bloecker H
MIPS

An Klopferspitz 18a D-82152 Martinsried, Germany

This is the 3' sequence of the clone insert
 Clone from S. Wiemann, Molecular Genome Analysis, German Cancer
 Research Center (DKFZ); Email s.wiemann@dkfz-heidelberg.de;
 sequenced by GBF (National Research Centre for Biotechnology Ltd.,
 Braunschweig/Germany) within the cDNA sequencing consortium of the
 German Genome Project.
 No r1 sequence available.
 This clone (DKFZp434P0828) is available at the RZPD in Berlin.
 Please contact the RZPD: Ressourcenzentrum, Heubnerweg 6, 14059
 Berlin-Charlottenburg, GERMANY; Email: clone@rzpd.de.

FEATURES
 source
 1. .571
 Location/Qualifiers
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /clone="DKFZp434P0828"
 /clone_lib="434 (synonym: htes3)"
 /tissue_type="testis"
 /dev_stage="adult"
 /lab_host="DH10B"
 /note="Vector; pSport1; Site_1: NotI; Site_2: SalI"
 BASE COUNT 122 a 168 c 149 g 132 t
 ORIGIN

Query Match 31.6%; Score 449.8; DB 10; Length 571;
 Best Local Similarity 99.6%; Pred. No. 3.4e-116;
 Matches 451; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
 QY 712 GTGTAGCCGTAGTTAGGCGACACCTTCAGAACTCTGTAGCAGCGCTACATACCTCGCT 771
 Db 463 GTGTAGCCGTAGTTAGGCGACACCTTCAGAACTCTGTAGCAGCGCTACATACCTCGCT 404
 QY 772 CTGCTAATCTGTTACACGTGCTGCTGCAGTGGCGGATAAGTCGTTCTTACCGGGTTG 831
 Db 403 CTGCTAATCTGTTACACGTGCTGCTGCAGTGGCGGATAAGTCGTTCTTACCGGGTTG 344
 QY 832 GACTCAAGACGATAGTTACCGGATAAGCGCGAGCGTTCGGGCTGAACGGGGGTTTCGTGC 891
 Db 343 GACTCAAGACGATAGTTACCGGATAAGCGCGAGCGTTCGGGCTGAACGGGGGTTTCGTGC 284
 QY 892 ACACAGCCAGCTTGGAGCGACGACCTACACGGAAGTACGATACCTACAGCGTGAGCAT 951
 Db 283 ACACAGCCAGCTTGGAGCGACGACCTACACGGAAGTACGATACCTACAGCGTGAGCTA 224
 QY 952 TGAGAAAGCGCCAGCTTCCGGAAGGAGAAAGCGGCGACAGGTATCCGGTAAAGCGGCGAG 1011
 Db 223 TGAGAAAGCGCCAGCTTCCGGAAGGAGAAAGCGGCGACAGGTATCCGGTAAAGCGGCGAG 164
 QY 1012 GTGGAAACAGGAGCGCGACGAGGGAGCTTCCAGGGGGAACGCTGATCTTTATAGT 1071
 Db 163 GTGGAAACAGGAGCGCGACGAGGGAGCTTCCAGGGGGAACGCTGATCTTTATAGT 104
 QY 1072 CCTGTCCGGTTTCGCCACCTCTGACTTGTAGCGTCGATTTTGTGATGCTCGTCAGGGGG 1131
 Db 103 CCTGTCCGGTTTCGCCACCTCTGACTTGTAGCGTCGATTTTGTGATGCTCGTCAGGGGG 44
 QY 1132 CGGAGCCTATGGAAGAAACGCCAGCAACCGCGCC 1164
 Db 43 CGGAGCCTATGGAAGAAACGCCAGCAACCGCGCC 11

RESULT 4
 AL043613/c
 LOCUS
 DEFINITION DKFZp434H1527_s1 434 (synonym: htes3) Homo sapiens cDNA clone
 ACCESSION. AL043613
 VERSION AL043613.1 GI:5423000
 KEYWORDS EST..
 SOURCE human.
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE
 AUTHORS
 TITLE
 JOURNAL
 COMMENT

1 (bases 1 to 579)
 Bloecker, H., Boecker, M., Brandt, P., Mewes, H.W., Gassenhuber, J. and
 Wiemann, S.
 EST (Bloecker, et al.)
 Unpublished (1999)
 Contact: Bloecker H
 MIPS
 Am Klopferspitz 18a D-82152 Martinsried, Germany
 This is the 3' sequence of the clone insert
 Clone from S. Wiemann, Molecular Genome Analysis, German Cancer
 Research Center (DKFZ); Email s.wiemann@dkfz-heidelberg.de;
 sequenced by GBF (National Research Centre for Biotechnology Ltd.,
 Braunschweig/Germany) within the cDNA sequencing consortium of the
 German Genome Project.
 No r1 sequence available.
 This clone (DKFZp434H1527) is available at the RZPD in Berlin.
 Please contact the RZPD: Ressourcenzentrum, Heubnerweg 6, 14059
 Berlin-Charlottenburg, GERMANY; Email: clone@rzpd.de.

FEATURES
 source
 1. .579
 Location/Qualifiers

/organism="Homo sapiens"
 /db_xref="taxon:9606"
 /clone="DKFZp434H1527"
 /clone_lib="434 (synonym: htes3)"
 /tissue_type="testis"
 /dev_stage="adult"
 /lab_host="DH10B"
 /note="Vector; pSport1; Site_1: NotI; Site_2: SalI"

BASE COUNT 128 a 171 c 152 g 128 t
 ORIGIN

Query Match 31.6%; Score 449.8; DB 10; Length 579;
 Best Local Similarity 99.6%; Pred. No. 3.4e-116;
 Matches 451; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
 QY 712 GTGTAGCCGTAGTTAGGCGACACCTTCAGAACTCTGTAGCAGCGCTACATACCTCGCT 771
 Db 477 GTGTAGCCGTAGTTAGGCGACACCTTCAGAACTCTGTAGCAGCGCTACATACCTCGCT 418
 QY 772 CTGCTAATCTGTTACACGTGCTGCTGCAGTGGCGGATAAGTCGTTCTTACCGGGTTG 831
 Db 417 CTGCTAATCTGTTACACGTGCTGCTGCAGTGGCGGATAAGTCGTTCTTACCGGGTTG 358
 QY 832 GACTCAAGACGATAGTTACCGGATAAGCGCGAGCGTTCGGGCTGAACGGGGGTTTCGTGC 891
 Db 357 GACTCAAGACGATAGTTACCGGATAAGCGCGAGCGTTCGGGCTGAACGGGGGTTTCGTGC 298
 QY 892 ACACAGCCAGCTTGGAGCGACGACCTACACCGAACTGAGATACCTACAGCGTGAGCAT 951
 Db 297 ACACAGCCAGCTTGGAGCGACGACCTACACCGAACTGAGATACCTACAGCGTGAGCTA 238
 QY 952 TGAGAAAGCGCCAGCTTCCGGAAGGAGAAAGCGGCGACAGGTATCCGGTAAAGCGGCGAG 1011
 Db 237 TGAGAAAGCGCCAGCTTCCGGAAGGAGAAAGCGGCGACAGGTATCCGGTAAAGCGGCGAG 178
 QY 1012 GTCGGAACAGGAGCGCGACGAGGAGCTTCCAGGGGGAACGCTGATCTTTATAGT 1071
 Db 177 GTCGGAACAGGAGCGCGACGAGGAGCTTCCAGGGGGAACGCTGATCTTTATAGT 118
 QY 1072 CTTGTCCGGTTTCGCCACCTCTGACTTGTAGCGTCGATTTTGTGATGCTCGTCAGGGGG 1131
 Db 117 CTTGTCCGGTTTCGCCACCTCTGACTTGTAGCGTCGATTTTGTGATGCTCGTCAGGGGG 58
 QY 1132 CGGAGCCTATGGAAGAAACGCCAGCAACCGCGCC 1164
 Db 57 CGGAGCCTATGGAAGAAACGCCAGCAACCGCGCC 25

RESULT 5
 AV735756/c
 LOCUS
 DEFINITION AV735756 CB Homo sapiens cDNA clone CBMAGC03 5', mRNA sequence.
 AV735756 616 bp mRNA EST 17-OCT-2000

ACCESSION AV735756
 VERSION AV735756.1 GI:10853337
 KEYWORDS EST.
 SOURCE human.
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 REFERENCE 1 (bases 1 to 616)
 AUTHORS Zhang, O., Ye, M., Wu, X., Gu, J., Huang, Q., Zhou, J., Shen, Y., Han, Z.,
 Chen, S., Mao, M., and Chen, Z.
 TITLE Homo sapiens CB library cDNA clones
 JOURNAL Unpublished (2000)
 COMMENT Contact: Zhu Chen
 Shanghai Institute of Hematology, Rui-Jin Hospital
 197 Rui-Jin II Road, Shanghai 200025, P. R. China
 Tel: 86-21-64740490
 Fax: 86-21-64743206
 Email: mbshl@ms.stn.sh.cn
 This clone is available at Shanghai Hematology Institute in
 Shanghai.
 Chinese National Human Genome Center at Shanghai
 351 Guo Shoujing Road, Zhangjiang Hi-Tech Park, Pudong.
 FEATURES
 source
 1. 616
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /clone="CBMAGC03"
 /clone_lib="CB"
 /tissue_type="cord blood"
 /cell_type="CD34+ hematopoietic stem/progenitor cell"
 /lab_host="BM25.8"
 /note="Vector: pBluescript; Site_1: EcoRI; The insert is
 cloned randomly with the EcoRI digestion"
 BASE COUNT 125 a 182 c 165 g 143 t
 ORIGIN
 Query Match 31.6%; Score 449.8; DB 10; Length 616;
 Best Local Similarity 99.6%; Pred. No. 3.5e-116;
 Matches 451; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
 QY 712 GTGTAGCGTAGTTAGGCCACCTTCAAGAACTCTGTAGCAGCGCTACATACCTCGCT 771
 Db 487 GTGTAGCGTAGTTAGGCCACCTTCAAGAACTCTGTAGCAGCGCTACATACCTCGCT 428
 QY 772 CTGCTAATCTGTACCAGTGGCTGTGCGAGTGGCGATAAGTCTGTCTTACCGGGTTG 831
 Db 427 CTGCTAATCTGTACCAGTGGCTGTGCGAGTGGCGATAAGTCTGTCTTACCGGGTTG 368
 QY 832 GACTCAAGACGATAGTTACCGGATAAGCGCAGCGGTGCGGTGAACGGGGTTCGTGC 891
 Db 367 GACTCAAGACGATAGTTACCGGATAAGCGCAGCGGTGCGGTGAACGGGGTTCGTGC 308
 QY 892 ACACAGCCAGCTTGAGCGACGACCTACACGAACTGAGATACCTACACGCTGACAT 951
 Db 307 ACACAGCCAGCTTGAGCGACGACCTACACGAACTGAGATACCTACACGCTGACAT 248
 QY 952 TGAGAAAGCCGACGCTTCCGAGAGGAGAAAGCGGACGATTCGGTAAGCGGCGAG 1011
 Db 247 TGAGAAAGCCGACGCTTCCGAGAGGAGAAAGCGGACGATTCGGTAAGCGGCGAG 188
 QY 1012 GTCCGAAAGGAGAGCGACGAGGAGCTTCCAGGGGAAACGCGTGTATCTTTAGT 1071
 Db 187 GTCCGAAAGGAGAGCGACGAGGAGCTTCCAGGGGAAACGCGTGTATCTTTAGT 128
 QY 1072 CCGTGTGGGTTTCGCGACCTCTGAGCTTACGGTTCGATTTTGTGATGCTCGTACGGGGG 1131
 Db 127 CCGTGTGGGTTTCGCGACCTCTGAGCTTACGGTTCGATTTTGTGATGCTCGTACGGGGG 68
 QY 1132 CGGAGCTATGGAAGGAGCGACGAGCGGCC 1164
 Db 67 CGGAGCTATGGAAGGAGCGACGAGCGGCC 35

RESULT 6
 AJ281661/c
 LOCUS
 DEFINITION 4A3A-P8G10-F Anopheles gambiae immune competent 4A3A Anopheles
 gambiae cDNA clone 4A3A-P8G10, mRNA sequence.
 ACCESSION AJ281661
 VERSION AJ281661.1 GI:6929540
 KEYWORDS EST.
 SOURCE African malaria mosquito.
 ORGANISM Anopheles gambiae
 Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
 Pterygota; Neoptera; Endopterygota; Diptera; Nematocera; Culicoidea
 ; Anopheles.
 REFERENCE 1 (bases 1 to 617)
 AUTHORS Dimopoulos, G., Casavant, T.L., Chang, S., Scheetz, T., Roberts, C.,
 Donohue, M., Schultz, J., Benes, V., Bork, P., Ansoorge, W., Soares, M.B.,
 and Kafatos, F.C.
 TITLE Anopheles gambiae pilot gene discovery project: identification of
 mosquito innate immunity genes from expressed sequence tags
 generated from immune-competent cell lines
 JOURNAL Proc. Natl. Acad. Sci. U.S.A. 97 (12), 6619-6624 (2000)
 MEDLINE 20300950
 COMMENT Contact: Dimopoulos G
 Fotis C. Kafatos Laboratory
 European Molecular Biology Laboratory
 Meyerhofstrasse 1, 69117 Heidelberg, Germany.
 FEATURES
 source
 1. 617
 /organism="Anopheles gambiae"
 /strain="4A r/r"
 /db_xref="taxon:7165"
 /clone="4A3A-P8G10"
 /clone_lib="Anopheles gambiae immune competent 4A3A"
 /cell_line="immune competent 4A3A"
 /lab_host="F. coli DH10B"
 /note="Vector: pT73D-Pac (Pharmacia) with a modified
 polylinker; Site_1: EcoRI; Site_2: NotI; sequenced from
 forward priming site which reads from the 3' end of the
 cDNA. The 4A3A is a directionally cloned and normalized
 cDNA library that was constructed from the 4A3A cell line
 oligo-T primed cDNA according to: Bonaldo, Lennon & Soares
 (1996). Normalization and Subtraction: Two approaches To
 Facilitate Gene Discovery, Genome Research 6, 791-806."
 BASE COUNT 136 a 175 c 162 g 144 t
 ORIGIN

Query Match 31.6%; Score 449.8; DB 10; Length 617;
 Best Local Similarity 99.6%; Pred. No. 3.5e-116;
 Matches 451; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
 QY 712 GTGTAGCGTAGTTAGGCCACCTTCAAGAACTCTGTAGCAGCGCTACATACCTCGCT 771
 Db 488 GTGTAGCGTAGTTAGGCCACCTTCAAGAACTCTGTAGCAGCGCTACATACCTCGCT 429
 QY 772 CTGCTAATCTGTACCAGTGGCTGTGCGAGTGGCGATAAGTCTGTCTTACCGGGTTG 831
 Db 428 CTGCTAATCTGTACCAGTGGCTGTGCGAGTGGCGATAAGTCTGTCTTACCGGGTTG 369
 QY 832 GACTCAAGACGATAGTTACCGGATAAGCGCAGCGGTGCGGTGAACGGGGTTCGTGC 891
 Db 368 GACTCAAGACGATAGTTACCGGATAAGCGCAGCGGTGCGGTGAACGGGGTTCGTGC 309
 QY 892 ACACAGCCAGCTTGAGCGACGACCTACACGAACTGAGATACCTACACGCTGACAT 951
 Db 308 ACACAGCCAGCTTGAGCGACGACCTACACGAACTGAGATACCTACACGCTGACAT 249
 QY 952 TGAGAAAGCCGACGCTTCCGAGAGGAGAAAGCGGACGATTCGGTAAGCGGCGAG 1011
 Db 248 TGAGAAAGCCGACGCTTCCGAGAGGAGAAAGCGGACGATTCGGTAAGCGGCGAG 189
 QY 1012 GTCCGAAAGGAGAGCGACGAGGAGCTTCCAGGGGAAACGCGTGTATCTTTAGT 1071

```
|||||
Db 188 GTGGAAACAGGAGCGCACGAGGAGCTTCCAGGGGAGAAACGCTGGTATCTTTATAGT 129
|||||
QY 1072 CCTGTCGGGTTTCGCACCTCTGACTTGAGCTCGATTTTGTGATGCTCGTCAGGGGG 1131
|||||
Db 128 CCTGTCGGGTTTCGCACCTCTGACTTGAGCTCGATTTTGTGATGCTCGTCAGGGGG 69
|||||
QY 1132 CGGAGCCTATGGAAGAACGCGCAGCAACGCGGCC 1164
|||||
Db 68 CGGAGCCTATGGAAGAACGCGCAGCAACGCGGCC 36
|||||

RESULT 7
BF381364/c 628 bp mRNA EST 27-NOV-2000
LOCUS
DEFINITION AsIR0004 Mosquito Genes Pool Related Malaria Infection Anopheles
stephensi cDNA 5', mRNA sequence.
ACCESSION BF381364
VERSION
KEYWORDS EST.
SOURCE Anopheles stephensi.
ORGANISM Anopheles stephensi.
REFERENCE Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
AUTHORS Pterygota; Neoptera; Endopterygota; Diptera; Nematocera; Culicoidea
; Anopheles.
TITLE 1 (bases 1 to 628)
XU,X., QU,F., SONG,G. and XU,J.
JOURNAL The differentially expressing genes pool from Anopheles stephensi
COMMENT related to infection with Plasmodium yoelii enriched by suppression
subtractive hybridization
Unpublished (2001)
Contact: Xu Xiaochun; Qu Fengyi; Song Guanhong; Xu Jiannong
Department of Parasitology
Second Military Medical University
800 Xiangyin Rd., Shanghai, 200433, China
Tel: 86 021 25070276
Email: xcxu@smmu.edu.cn
Seq primer: T7
High quality sequence stop: 629
POLYA=No. Location/Qualifiers
source 1..628
/organism="Anopheles stephensi"
/strain="Hor"
/db_xref="taxon:30069"
/clone_lib="Mosquito Genes Pool Related Malaria Infection"
/sex="female"
/tissue_type="whole body"
/dev_stage="24 hours post-infection"
BASE COUNT 145 a 178 c 169 g 136 t
ORIGIN

Query Match 31.6%; Score 449.8; DB 11; Length 628;
Best Local Similarity 99.6%; Pred. No. 3.5e-116;
Matches 451; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 712 GTGTAGCCGTAGTTAGCCACCCTTCAAGAACTCTGTAGCAGCGCTGAAACGGGGTTCGTGC 771
|||||
Db 569 GTGTAGCCGTAGTTAGCCACCCTTCAAGAACTCTGTAGCAGCGCTGAAACGGGGTTCGTGC 510
|||||
QY 772 CTGCTAATCTGTTTACCAGTGGCTGTCGCCAGTGGCGGATAAGTCTCTTACCGGGTTG 831
|||||
Db 509 CTGCTAATCTGTTTACCAGTGGCTGTCGCCAGTGGCGGATAAGTCTCTTACCGGGTTG 450
|||||
QY 832 GACTCAAGACGATAGTTACCGGATAAGCGCAGCGGTTCGGGCTGAAACGGGGTTCGTGC 891
|||||
Db 449 GACTCAAGACGATAGTTACCGGATAAGCGCAGCGGTTCGGGCTGAAACGGGGTTCGTGC 390
|||||
QY 892 ACACAGCCCAGCTTGGAGCGAAGACCTACACCGAACTGAGATACCTACACGCGGTGACAT 951
|||||
Db 389 ACACAGCCCAGCTTGGAGCGAAGACCTACACCGAACTGAGATACCTACACGCGGTGACAT 930
|||||
```

```
QY 952 TGAGAAAGCCGCCACGCTTCCGAAAGGAGGAGAAAGCGGACAGGTATCCGTTAGCGGCAGG 1011
|||||
Db 329 TGAGAAAGCCGCCACGCTTCCGAAAGGAGGAGAAAGCGGACAGGTATCCGTTAGCGGCAGG 270
|||||
QY 1012 GTGGAACAGGAGAGCGCACGAGGAGCTTCCAGGGGAGAAACGCTGTTATCTTTATAGT 1071
|||||
Db 269 GTGGAACAGGAGAGCGCACGAGGAGCTTCCAGGGGAGAAACGCTGTTATCTTTATAGT 210
|||||
QY 1072 CCTGTCGGGTTTCGCCACCTCTGACTTGAGCTCGATTTTGTGATGCTCGTCAGGGGG 1131
|||||
Db 209 CCTGTCGGGTTTCGCCACCTCTGACTTGAGCTCGATTTTGTGATGCTCGTCAGGGGG 150
|||||
QY 1132 CGGAGCCTATGGAAGAACGCGCAGCAACGCGGCC 1164
|||||
Db 149 CGGAGCCTATGGAAGAACGCGCAGCAACGCGGCC 117
|||||

RESULT 8
AL044364/c 954 bp mRNA EST 29-FEB-2000
LOCUS
DEFINITION DKFZp434C172.sl 434 (synonym: htes3) Homo sapiens cDNA clone
ACCESSION AL044364
VERSION DKFZp434C172.3', mRNA sequence.
KEYWORDS EST.
SOURCE AL044364.1 GI:5432586
human.
ORGANISM Homo sapiens
REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
AUTHORS Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
Wiemann,S.
Ansorge,W., Benes,V., Krieger,S., Mewes,H.W., Gassenhuber,J. and
EST (Ansorge, Benes, et al.)
Unpublished (1999)
Contact: Ansorge W
MIPS
TITLE Am Klopferspitz 18a D-82152 Martinsried, Germany
JOURNAL This is the 3' sequence of the clone insert
COMMENT Clone from S. Wiemann, Molecular Genome Analysis, German Cancer
Research Center (DKFZ); Email s.wiemann@dkfz-heidelberg.de;
Heidelberg/Germany) within the cDNA sequencing consortium of the
German Genome Project.
No r1 sequence available.
This clone (DKFZp434C172) is available at the RZPD in Berlin.
Please contact the RZPD: Ressourcenzentrum, Heubnerweg 6, 14059
Berlin-Charlottenburg, GERMANY; Email: clone@rzpd.de.
Location/Qualifiers
source 1..954
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="DKFZp434C172"
/clone_lib="434 (synonym: htes3)"
/tissue_type="testis"
/dev_stage="adult"
/lab_host="DH10B"
/notes="Vector: pSport1; Site_1: NotI; Site_2: SalI"
BASE COUNT 214 a 273 c 253 g 214 t
ORIGIN

Query Match 31.6%; Score 449.8; DB 10; Length 954;
Best Local Similarity 99.6%; Pred. No. 4e-116;
Matches 451; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 712 GTGTAGCCGTAGTTAGCCACCCTTCAAGAACTCTGTAGCAGCGCTTACATACCTCGCT 771
|||||
Db 851 GTGTAGCCGTAGTTAGCCACCCTTCAAGAACTCTGTAGCAGCGCTTACATACCTCGCT 792
|||||
QY 772 CTGCTAATCTGTTTACCAGTGGCTGTCGCCAGTGGCGGATAAGTCTGTCTTACCGGGTTG 831
|||||
Db 791 CTGCTAATCTGTTTACCAGTGGCTGTCGCCAGTGGCGGATAAGTCTGTCTTACCGGGTTG 732
|||||
```

```

QY 832 GACTCAAGCAGCATAGTTACCGGATAAGCGCAGCGTTCGGCTGAACGGGGGGTTTCGTGC 891
Db 731 GACTCAAGCAGCATAGTTACCGGATAAGCGCAGCGTTCGGCTGAACGGGGGGTTTCGTGC 672
QY 892 ACACAGCCAGCTTTGGAGCGAAGCAGCTACACCACTGAGATACCTACAGCGTGAACAT 951
Db 671 ACACAGCCAGCTTTGGAGCGAAGCAGCTACACCACTGAGATACCTACAGCGTGAACAT 612
QY 952 TGAGAAAGCGCCAGCTTCCGAGGAGAGAAAGCGGACAGGTATCCGTAAGCGGCAGG 1011
Db 611 TGAGAAAGCGCCAGCTTCCGAGGAGAGAAAGCGGACAGGTATCCGTAAGCGGCAGG 552
QY 1012 GTCGGAACAGGAGCGCAGGAGCTTCCAGGGGAAACCGCTGATCTTTATAGT 1071
Db 551 GTCGGAACAGGAGCGCAGGAGCTTCCAGGGGAAACCGCTGATCTTTATAGT 492
QY 1072 CTTGTGGGTTTCCGACCTCTGACTTGAGCGTCGATTTTCTGATGCTGCTCAGGGGGG 1131
Db 491 CTTGTGGGTTTCCGACCTCTGACTTGAGCGTCGATTTTCTGATGCTGCTCAGGGGGG 432
QY 1132 CGGAGCTATGAAAGACCGCAGCAGCGGCC 1164
Db 431 CGGAGCTATGAAAGACCGCAGCAGCGGCC 399

RESULT 9
AL593919/c 629 bp mRNA EST 30-JUL-2001
LOCUS
DEFINITION
AL593919 XGC-gastrula silurana tropicalis cDNA clone TGas003010 5',
mRNA sequence.
ACCESSION
AL593919
VERSION
AL593919.1 GI:15005980
KEYWORDS
EST.
SOURCE
western clawed frog.
ORGANISM
Silurana tropicalis
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Amphibia; Batrachia; Anura; Mesobatrachia; Pipoidae; Pipidae;
Xenopodinae; Silurana.
1 (bases 1 to 629)
Huckle, E., Taylor, R., McMurray, A., Ashurst, J. L., Zorn, A. M. and
Rogers, J.
Sanger Xenopus tropicalis EST project 2001
Unpublished (2001)
Contact: Huckle E
Sanger Centre
Hinxtion, Cambridgeshire, CB10 1SA, UK
Email: trop@sanger.ac.uk
Sanger Xenopus tropicalis EST project 2001
TROPICALIS_SEQUENCE_ID: TGas003010.sp6
Sequencing primer: SP6
This sequence is from a Xenopus Gene Collection (XGC) library
constructed by Aaron M. Zorn.
FEATURES
Location/Qualifiers
1..629
/organism="Silurana tropicalis"
/db_xref="taxon:8364"
/clone="TGas003010"
/clone_lib="XGC-gastrula"
/dev_stage="gastrula (stages 10.5-13 mixed)"
/lab_host="Escherichia coli DH10B"
/note="Vector: pCS107; Site_1: EcoRI; Site_2: NotI; cDNA
was oligo dt primed from 5ug of poly A+ RNA from stages
10-13 gastrulae. EcoRI-NotI cut cDNA was then ligated
into pCS107 with EcoRI at the 5' end and NotI at the 3'
end."
BASE COUNT 149 a 181 c 164 g 135 t
ORIGIN

Query Match 31.5%; Score 448.8; DB 10; Length 629;
Best Local Similarity 99.6%; Pred. No. 6.7e-116;
Matches 450; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

```

```

QY 712 GTGTAGCGGTAGTTAGCGCACCCTTCAAGAACTCTGTAGCACCCTACATACCTCGCT 771
Db 528 GTGTAGCGGTAGTTAGCGCACCCTTCAAGAACTCTGTAGCACCCTACATACCTCGCT 469
QY 772 CTGCTAATCTGTTACCAAGTGGCTGCTGCCAGTGGCGATAAGTCGTCTTACCGGGTTG 831
Db 468 CTGCTAATCTGTTACCAAGTGGCTGCTGCCAGTGGCGATAAGTCGTCTTACCGGGTTG 409
QY 832 GACTCAAGCAGTATGTTACCGGATAAGCGCAGCGTTCGGCTGAACGGGGGGTTTCGTGC 891
Db 408 GACTCAAGCAGTATGTTACCGGATAAGCGCAGCGTTCGGCTGAACGGGGGGTTTCGTGC 349
QY 892 ACACAGCCAGCTTGGAGCGAAGCAGCTACACCACTGAGATACCTACAGCGTGAACAT 951
Db 348 ACACAGCCAGCTTGGAGCGAAGCAGCTACACCACTGAGATACCTACAGCGTGAACAT 289
QY 952 TGAGAAAGCGCCAGCTTCCGAGGAGAGAAAGCGGACAGGTATCCGTAAGCGGCAGG 1011
Db 288 TGAGAAAGCGCCAGCTTCCGAGGAGAGAAAGCGGACAGGTATCCGTAAGCGGCAGG 229
QY 1012 GTCGGAACAGGAGCGCAGGAGCTTCCAGGGGAAACCGCTGATCTTTATAGT 1071
Db 228 GTCGGAACAGGAGCGCAGGAGCTTCCAGGGGAAACCGCTGATCTTTATAGT 169
QY 1072 CTTGTGGGTTTCCGACCTCTGACTTGAGCGTCGATTTTCTGATGCTGCTCAGGGGGG 1131
Db 168 CTTGTGGGTTTCCGACCTCTGACTTGAGCGTCGATTTTCTGATGCTGCTCAGGGGGG 109
QY 1132 CGGAGCTATGAAAGACCGCAGCAGCGGCC 1163
Db 108 CGGAGCTATGAAAGACCGCAGCAGCGGCC 77

RESULT 10
AL593919/c 1004 bp mRNA EST 30-JUN-2000
LOCUS
DEFINITION
AL593919 XGC-gastrula silurana tropicalis cDNA clone TGas003010 5',
mRNA sequence.
ACCESSION
AL593919
VERSION
AL593919.1 GI:15005980
KEYWORDS
EST.
SOURCE
African malaria mosquito.
ORGANISM
Anopheles gambiae
Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
Pterygota; Neoptera; Endopterygota; Diptera; Nematocera; Culicoidea
; Anopheles.
1 (bases 1 to 1004)
Dimopoulos, G., Casavant, T. L., Chang, S., Scheetz, T., Roberts, C.,
Donohue, M., Schultz, J., Benes, V., Bork, P., Ansorge, W., Soares, M. B.
and Kafatos, F. C.
Anopheles gambiae pilot gene discovery project: identification of
mosquito innate immunity genes from expressed sequence tags
generated from immune-competent cell lines
Proc. Natl. Acad. Sci. U.S.A. 97 (12), 6619-6624 (2000)
20300950
Contact: Dimopoulos G
Fotis C. Kafatos laboratory
European Molecular Biology Laboratory
Meyhofstrasse 1, 69117 Heidelberg, Germany.
Location/Qualifiers
1..1004
/organism="Anopheles gambiae"
/db_xref="taxon:7165"
/clone="4A3A-P4G8"
/clone_lib="Anopheles gambiae immune competent 4A3A"
/cell_line="E. coli DH10B"
/lab_host="E. coli DH10B"
/note="Vector: pT73D-Pac (Pharmacia) with a modified
polylinker; Site_1: EcoRI; Site_2: NotI; sequenced from
forward priming site which reads from the 3' end of the
cDNA. The 4A3A is a directionally cloned and normalized
cDNA library that was constructed from the 4A3A cell line

```

oligo-T primed cDNA according to: Bonaldo, Lennon & Soares (1996) : Normalization and Subtraction: Two approaches to Facilitate Gene Discovery, Genome Research 6, 791-806."

BASE COUNT 252 a 262 c 244 g 244 t 2 others

ORIGIN

Query Match 31.5%; Score 448.8; DB 10; Length 1004;
Best Local Similarity 99.3%; Pred. No. 7.8e-116;
Matches 450; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 712 GTGTAGCGTAGTTAGCCACCACTTCAAGAACTCTGTAGCAGCGCTACATACCTCGCT 771
DB 475 GTGTAGCGTAGTTAGCCACCACTTCAAGAACTCTGTAGCAGCGCTACATACCTCGCT 416
QY 772 CTGCTAATCTGTTACCAAGTGGCTGCTGCCAGTGGCGATAAGTCTCTTACCGGGTTG 831
DB 415 CTGCTAATCTGTTACCAAGTGGCTGCTGCCAGTGGCGATAAGTCTCTTACCGGGTTG 356
QY 832 GACTCAAGACGATGTTACCGGATAAGCGCAGCGTCCGGCTGAAGGGGGTTCGTGC 891
DB 355 GACTCAAGACGATGTTACCGGATAAGCGCAGCGTCCGGCTGAAGGGGGTTCGTGC 296
QY 892 ACACAGCCCGAGTTGAGCGAAGACCTACACGAACTGAGATACCTACAGCGTGAGCAT 951
DB 295 ANACAGCCCGAGTTGAGCGAAGACCTACACGAACTGAGATACCTACAGCGTGAGCTA 236
QY 952 TGAGAAAGCCCGAGCTTCCCGAAGGAGGAGAAAGCGGACAGGTATCCGGTAAGCGGCAGG 1011
DB 235 TGAGAAAGCCCGAGCTTCCCGAAGGAGGAGAAAGCGGACAGGTATCCGGTAAGCGGCAGG 176
QY 1012 GTCGGAACAGGAGCGCAGGAGGAGCTTCCAGGGGAAACGCTGGTATCTTTATAGT 1071
DB 175 GTCGGAACAGGAGCGCAGGAGGAGCTTCCAGGGGAAACGCTGGTATCTTTATAGT 116
QY 1072 CCGTGGGTTTCGCCACCTCTGACTTGAGCGTCGATTTTGTGATGCTCGTCAGGGGG 1131
DB 115 CCGTGGGTTTCGCCACCTCTGACTTGAGCGTCGATTTTGTGATGCTCGTCAGGGGG 56
QY 1132 CGGAGCCTATGGAAGAAACGCCAGCGCGCC 1164
DB 55 CGGAGCCTATGGAAGAAACGCCAGCGCGCC 23

RESULT 11
AJ281437/37
LOCUS 703 bp mRNA EST 30-JUN-2000
DEFINITION 4A3A-P4C3-F Anopheles gambiae immune competent 4A3A Anopheles gambiae cDNA clone 4A3A-P4C3, mRNA sequence.

ACCESSION AJ281437
VERSION AJ281437.1 GI:6929317
KEYWORDS EST.
SOURCE African malaria mosquito.
ORGANISM Anopheles gambiae

REFERENCE 1 (bases 1 to 703)
AUTHORS Dimopoulos G., Casavant, T.L., Chang, S., Scheetz, T., Roberts, C., Donohue, M., Schultz, J., Benes, V., Bork, P., Ansorge, W., Soares, M.B. and Kafatos, F.C.

TITLE mosquito innate immunity genes from expressed sequence tags generated from immune-competent cell lines
JOURNAL Proc. Natl. Acad. Sci. U.S.A. 97 (12), 6619-6624 (2000)
MEDLINE 20300950
COMMENT Contact: Dimopoulos G
European Molecular Biology Laboratory
Meyerhofstrasse 1, 69117 Heidelberg, Germany.
Location/Qualifiers
1. .703
/organism="Anopheles gambiae"

FEATURES
Source

/strain="4A r/r"
/db_xref="taxon:7165"
/clone="4A3A-P4C3"
/cell_line="immune competent 4A3A"
/lab_host="E. coli DH10B"
/note="Vector: pT73D-Pac (Pharmacia) with a modified polylinker; Site_1: EcoRI; Site_2: NotI; sequenced from forward priming site which reads from the 3' end of the cDNA. The 4A3A is a directionally cloned and normalized cDNA library that was constructed from the 4A3A cell line oligo-T primed cDNA according to: Bonaldo, Lennon & Soares (1996) : Normalization and Subtraction: Two approaches to Facilitate Gene Discovery, Genome Research 6, 791-806."
BASE COUNT 159 a 204 c 193 g 147 t
ORIGIN

Query Match 31.5%; Score 448.2; DB 10; Length 703;
Best Local Similarity 99.3%; Pred. No. 1e-115;
Matches 450; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 712 GTGTAGCGTAGTTAGCCACCACTTCAAGAACTCTGTAGCAGCGCTACATACCTCGCT 771
DB 692 GTGTAGCGTAGTTAGCCACCACTTCAAGAACTCTGTAGCAGCGCTACATACCTCGCT 633
QY 772 CTGCTAATCTGTTACCAAGTGGCTGCTGCCAGTGGCGATAAGTCTCTTACCGGGTTG 831
DB 632 CTGCTAATCTGTTACCAAGTGGCTGCTGCCAGTGGCGATAAGTCTCTTACCGGGTTG 573
QY 832 GACTCAAGACGATGTTACCGGATAAGCGCAGCGTCCGGCTGAAGGGGGTTCGTGC 891
DB 572 GACTCAAGACGATGTTACCGGATAAGCGCAGCGTCCGGCTGAAGGGGGTTCGTGC 513
QY 892 ACACAGCCCGAGTTGAGCGAAGACCTACACGAACTGAGATACCTACAGCGTGAGCAT 951
DB 512 ACACAGCCCGAGTTGAGCGAAGACCTACACGAACTGAGATACCTACAGCGTGAGCTA 453
QY 952 TGAGAAAGCCCGAGCTTCCCGAAGGAGGAGAAAGCGGACAGGTATCCGGTAAGCGGCAGG 1011
DB 452 TGAGAAAGCCCGAGCTTCCCGAAGGAGGAGAAAGCGGACAGGTATCCGGTAAGCGGCAGG 393
QY 1012 GTCGGAACAGGAGCGCAGGAGGAGCTTCCAGGGGAAACGCTGGTATCTTTATAGT 1071
DB 392 GTCGGAACAGGAGCGCAGGAGGAGCTTCCAGGGGAAACGCTGGTATCTTTATAGT 333
QY 1072 CCGTGGGTTTCGCCACCTCTGACTTGAGCGTCGATTTTGTGATGCTCGTCAGGGGG 1131
DB 332 CCGTGGGTTTCGCCACCTCTGACTTGAGCGTCGATTTTGTGATGCTCGTCAGGGGG 273
QY 1132 CGGAGCCTATGGAAGAAACGCCAGCGCGCC 1164
DB 272 CGGAGCCTATGGAAGAAACGCCAGCGCGCC 240

RESULT 12
AJ281449/49
LOCUS 800 bp mRNA EST 30-JUN-2000
DEFINITION 4A3A-P4D5-F Anopheles gambiae immune competent 4A3A Anopheles gambiae cDNA clone 4A3A-P4D5, mRNA sequence.

ACCESSION AJ281449
VERSION AJ281449.1 GI:6929329
KEYWORDS EST.
SOURCE African malaria mosquito.
ORGANISM Anopheles gambiae

REFERENCE 1 (bases 1 to 800)
AUTHORS Dimopoulos G., Casavant, T.L., Chang, S., Scheetz, T., Roberts, C., Donohue, M., Schultz, J., Benes, V., Bork, P., Ansorge, W., Soares, M.B. and Kafatos, F.C.

TITLE Anopheles gambiae pilot gene discovery project: identification of

mosquito innate immunity genes from expressed sequence tags
generated from immune-competent cell lines
Proc. Natl. Acad. Sci. U.S.A. 97 (12), 6619-6624 (2000)
20300950
Contact: Dimopoulos G
Fotis C. Kafatos laboratory
European Molecular Biology Laboratory
Meyerhofstrasse 1, 69117 Heidelberg, Germany.
Location/Qualifiers
1. .800
/organism="Anopheles gambiae"
/strain="4A r/r"
/db_xref="taxon:7165"
/clone="4A3A-P405"
/cell_line="Immune competent 4A3A"
/lab_host="E. coli DH10B"
/note="Vector: p773D-Pac (Pharmacia) with a modified
polylinker; Site_1: EcoRI; Site_2: NotI; sequenced from
forward priming site which reads from the 3' end of the
cDNA. The 4A3A is a directionally cloned and normalized
cDNA library that was constructed from the 4A3A cell line
oligo-T primed cDNA according to: Bonaldo, Lennon & Soares
(1996) : Normalization and Subtraction: Two approaches To
Facilitate Gene Discovery, Genome Research 6, 791-806."

BASE COUNT 203 a 206 c 198 g 193 t
ORIGIN

Query Match 31.5%; Score 448.2; DB 10; Length 800;
Best Local Similarity 99.3%; Pred. No. 1.1e-115;
Matches 450; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 712 GTGTAGCCGTAGTTAGCCACCACTTCAAGAACTCTGTAGCACCGCTACATACCTCGCT 771
|||||
Db 502 GTGTAGCCGTAGTTAGCCACCACTTCAAGAACTCTGTAGCACCGCTACATACCTCGCT 771
|||||
QY 772 CTGCTAATCTGTTACCACTGCTGCCAGTGGCGATAAGTCTGCTTACCGGGTTG 831
|||||
Db 442 CTGCTAATCTGTTACCACTGCTGCCAGTGGCGATAAGTCTGCTTACCGGGTTG 831
|||||
QY 832 GACTCAAGACGATAGTTACCGATAAGCGCAGCGTGGCGTGAACGGGGGTTCTGTC 891
|||||
Db 382 GACTCAAGACGATAGTTACCGATAAGCGCAGCGTGGCGTGAACGGGGGTTCTGTC 891
|||||
QY 892 ACACAGCCACCTTGGAGCGAAGCACTACACCGAAGTGAATACCTACAGCTGAGCAT 951
|||||
Db 322 ACACAGCCACCTTGGAGCGAAGCACTACACCGAAGTGAATACCTACAGCTGAGCAT 951
|||||
QY 952 TGAGAAAGCGCCACGCTTCCGGAAGGAGAGAGCGGAGGATCCGGTAAGCGGCAGG 1011
|||||
Db 262 TGAGAAAGCGCCACGCTTCCGGAAGGAGAGAGCGGAGGATCCGGTAAGCGGCAGG 1011
|||||
QY 1012 GTCGGAACGAGGAGCGGAGGAGCTTCCAGGGGAAACCGCTGTATCTTTATAGT 1071
|||||
Db 202 GTCGGAACGAGGAGCGGAGGAGCTTCCAGGGGAAACCGCTGTATCTTTATAGT 1071
|||||
QY 1072 CCGTCCGGTTTCGCCACCTCTGACTTGAGCTCGATTTTGTGATGCTCGTCAGGGGG 1131
|||||
Db 142 CCGTCCGGTTTCGCCACCTCTGACTTGAGCTCGATTTTGTGATGCTCGTCAGGGGG 1131
|||||
QY 1132 CGGAGCCTATGAAAAACGCCAGCAACCGGCC 1164
|||||
Db 82 CGGAGCCTATGAAAAACGCCAGCAACCGGCC 50
|||||

RESULT 13
AJ281376/c
LOCUS
DEFINITION 4A3A-P2G2-F Anopheles gambiae Immune competent 4A3A Anopheles
gambiae cDNA clone 4A3A-P2G2, mRNA sequence.
ACCESSION AJ281376
VERSION AJ281376.1 GI:6929257

EST.
SOURCE African malaria mosquito.
ORGANISM Anopheles gambiae
Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
Pterygota; Neoptera; Endopterygota; Diptera; Nematocera; Culicoidea
; Anopheles
1 (bases 1 to 568)
Dimopoulos, G., Casavant, T.L., Chang, S., Scheetz, T., Roberts, C.,
Donohue, M., Schultz, J., Benes, V., Bork, P., Ansorge, W., Soares, M.B.
and Kafatos, F.C.
Anopheles gambiae pilot gene discovery project: identification of
mosquito innate immunity genes from expressed sequence tags
generated from immune-competent cell lines
Proc. Natl. Acad. Sci. U.S.A. 97 (12), 6619-6624 (2000)
20300950
Contact: Dimopoulos G
Fotis C. Kafatos laboratory
European Molecular Biology Laboratory
Meyerhofstrasse 1, 69117 Heidelberg, Germany.
Location/Qualifiers
1. .568
/organism="Anopheles gambiae"
/strain="4A r/r"
/db_xref="taxon:7165"
/clone="4A3A-P2G2"
/cell_line="Immune competent 4A3A"
/lab_host="E. coli DH10B"
/note="Vector: p773D-Pac (Pharmacia) with a modified
polylinker; Site_1: EcoRI; Site_2: NotI; sequenced from
forward priming site which reads from the 3' end of the
cDNA. The 4A3A is a directionally cloned and normalized
cDNA library that was constructed from the 4A3A cell line
oligo-T primed cDNA according to: Bonaldo, Lennon & Soares
(1996) : Normalization and Subtraction: Two approaches To
Facilitate Gene Discovery, Genome Research 6, 791-806."

BASE COUNT 123 a 167 c 149 g 127 t
ORIGIN

Query Match 31.4%; Score 447.8; DB 10; Length 568;
Best Local Similarity 99.1%; Pred. No. 1.2e-115;
Matches 449; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 712 GTGTAGCCGTAGTTAGCCACCACTTCAAGAACTCTGTAGCACCGCTACATACCTCGCT 771
|||||
Db 466 GTGTAGCCGTAGTTAGCCACCACTTCAAGAACTCTGTAGCACCGCTACATACCTCGCT 771
|||||
QY 772 CTGCTAATCTGTTACCACTGCTGCCAGTGGCGATAAGTCTGCTTACCGGGTTG 831
|||||
Db 406 CTGCTAATCTGTTACCACTGCTGCCAGTGGCGATAAGTCTGCTTACCGGGTTG 831
|||||
QY 832 GACTCAAGACGATAGTTACCGATAAGCGCAGCGTGGCGTGAACGGGGGTTCTGTC 891
|||||
Db 346 GACTCAAGACGATAGTTACCGATAAGCGCAGCGTGGCGTGAACGGGGGTTCTGTC 891
|||||
QY 892 ACACAGCCACCTTGGAGCGAAGCACTACACCGAAGTGAATACCTACAGCTGAGCAT 951
|||||
Db 286 ACACAGCCACCTTGGAGCGAAGCACTACACCGAAGTGAATACCTACAGCTGAGCAT 951
|||||
QY 952 TGAGAAAGCGCCACGCTTCCGGAAGGAGAGAGCGGAGGATCCGGTAAGCGGCAGG 1011
|||||
Db 226 TGAGAAAGCGCCACGCTTCCGGAAGGAGAGAGCGGAGGATCCGGTAAGCGGCAGG 1011
|||||
QY 1012 GTCGGAACGAGGAGCGGAGGAGCTTCCAGGGGAAACCGCTGTATCTTTATAGT 1071
|||||
Db 166 GTCGGAACGAGGAGCGGAGGAGCTTCCAGGGGAAACCGCTGTATCTTTATAGT 1071
|||||
QY 1072 CCGTCCGGTTTCGCCACCTCTGACTTGAGCTCGATTTTGTGATGCTCGTCAGGGGG 1131
|||||
Db 106 CCGTCCGGTTTCGCCACCTCTGACTTGAGCTCGATTTTGTGATGCTCGTCAGGGGG 1131
|||||
QY 1132 CGGAGCCTATGAAAAACGCCAGCAACCGGCC 1164
|||||

KEYWORDS
SOURCE
ORGANISM
REFERENCE
AUTHORS
TITLE
JOURNAL
MEDLINE
COMMENT
FEATURES
source

```

|||||
Db 46 CGGAGCCTATGGAAGCGCCNCCAGCGGCC 14

RESULT 14
AJ281320/c
LOCUS
DEFINITION
4A3A-PIH1-F Anopheles gambiae immune competent 4A3A Anopheles
gambiae cDNA clone 4A3A-PIH1, mRNA sequence.
ACCESSION
AJ281320
VERSION
AJ281320.1 GI:6929201
KEYWORDS
EST.
SOURCE
African malaria mosquito.
ORGANISM
Anopheles gambiae
REFERENCE
1 (bases 1 to 579)
AUTHORS
Dimopoulos,G., Casavant,T.L., Chang,S., Scheetz,T., Roberts,C.,
Donohue,M., Schultz,J., Benes,V., Bork,P., Ansorge,W., Soares,M.B.
and Kafatos,F.C.
TITLE
Anopheles gambiae pilot gene discovery project: identification of
mosquito innate immunity genes from expressed sequence tags
generated from immune-competent cell lines
JOURNAL
Proc. Natl. Acad. Sci. U.S.A. 97 (12), 6619-6624 (2000)
MEDLINE
20300950
COMMENT
Contact: Dimopoulos G
Fotis C. Kafatos laboratory
European Molecular Biology Laboratory
Meyerothofstrasse 1, 69117 Heidelberg, Germany.
LOCATION/Qualifiers
1. .579
/organism="Anopheles gambiae"
/strain="4A r/r"
/db_xref="taxon:7165"
/clone_lib="4A3A-PIH1"
/clone_line="Anopheles gambiae immune competent 4A3A"
/cell_host="E. coli DH10B"
/notes="Vector: pT73D-Pac (Pharmacia) with a modified
polylinker; Site_1: EcoRI; Site_2: NotI; sequenced from
forward priming site which reads from the 3' end of the
cDNA. The 4A3A is a directionally cloned and normalized
cDNA library that was constructed from the 4A3A cell line
oligo-T primed cDNA according to: Bonaldo, Lennon & Soares
(1996) : Normalization and Subtraction: Two approaches To
Facilitate Gene Discovery, Genome Research 6, 791-806."
BASE COUNT 120 a 167 c 151 g 137 t 4 others
ORIGIN

Query Match 31.4%; Score 447.8; DB 10; Length 579;
Best Local Similarity 99.1%; Pred. No. 1.2e-115;
Matches 449; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 712 GTGTAGCGGTAGTGGCCACCACCTTCAAGAACTCTGTAGCAGCGCTACATACCTCGCT 771
|||||
Db 456 GTGTAGCGGTAGTGGCCACCACCTTCAAGAACTCTGTAGCAGCGCTACATACCTCGCT 397
|||||
QY 772 CTGCTAATCCTGTTACCAAGTGGCTGCTGCGAGTGGCGATAAGTCGTGTCTTACCGGGTTG 831
|||||
Db 396 CTGCTAATCCTGTTACCAAGTGGCTGCTGCGAGTGGCGATAAGTCGTGTCTTACCGGGTTG 337
|||||
QY 832 GACTCAAGCAGTATGTTACCGGATAGGCGCAGCGGTGCGGGTGTCTGTGTC 891
|||||
Db 336 GACTCAAGCAGTATGTTACCGGATAGGCGCAGCGGTGCGGGTGTCTGTGTC 277
|||||
QY 892 ACACAGCCAGCTTGGAGCGAAGCAGCTACACGAACTGAGATACCTACAGCGTGAGCAT 951
|||||
Db 276 ACACAGCCAGCTTGGAGCGAAGCAGCTACACGAACTGAGATACCTACAGCGTGAGCTA 217
|||||
QY 952 TGAGAAAGCGCCACGCTTCCCGAAGGAGAGAAAGCGGACAGGATATCCGGTAAAGCGGAGG 1011
|||||

```

```

Db 216 TGAGAAAGCGCCACGCTTCCGAAGGAGAGAAAGCGGACAGGTATCCGGTAAAGCGGAGG 157
QY 1012 GTCGGAACAGGAGAGCGCAGGAGGAGCTTCCAGGGGGAACCGCTCGTATCTTTATAGT 1071
|||||
Db 156 GTCGGAACAGGAGAGCGCAGGAGGAGCTTCCAGGGGGAACCGCTCGTATCTTTATAGT 97
QY 1072 CCGTGTGGGTTTCCGCCACCTCTGACTTGAGGTCGCTGATTTTGTGATCTGCTCAGGGGGG 1131
|||||
Db 96 CCGTGTGGGTTTCCGCCACCTCTGASGTGAGGTCGATTTTGTGATGCTCGTCAGGGGGG 37
QY 1132 CGGAGCCTATGGAAGCGCCAGCAACGCGGCC 1164
|||||
Db 36 CGGAGCCTATGGAAGCGCCAGCAACGCGGCC 4

RESULT 15
AJ281616/c
LOCUS
DEFINITION
4A3A-P8A12-F Anopheles gambiae immune competent 4A3A Anopheles
gambiae cDNA clone 4A3A-P8A12, mRNA sequence.
ACCESSION
AJ281616
VERSION
AJ281616.1 GI:6929495
KEYWORDS
EST.
SOURCE
African malaria mosquito.
ORGANISM
Anopheles gambiae
REFERENCE
1 (bases 1 to 700)
AUTHORS
Dimopoulos,G., Casavant,T.L., Chang,S., Scheetz,T., Roberts,C.,
Donohue,M., Schultz,J., Benes,V., Bork,P., Ansorge,W., Soares,M.B.
and Kafatos,F.C.
TITLE
Anopheles gambiae pilot gene discovery project: identification of
mosquito innate immunity genes from expressed sequence tags
generated from immune-competent cell lines
JOURNAL
Proc. Natl. Acad. Sci. U.S.A. 97 (12), 6619-6624 (2000)
MEDLINE
20300950
COMMENT
Contact: Dimopoulos G
Fotis C. Kafatos laboratory
European Molecular Biology Laboratory
Meyerothofstrasse 1, 69117 Heidelberg, Germany.
LOCATION/Qualifiers
1. .700
/organism="Anopheles gambiae"
/strain="4A r/r"
/db_xref="taxon:7165"
/clone_lib="4A3A-P8A12"
/clone_line="Anopheles gambiae immune competent 4A3A"
/cell_host="E. coli DH10B"
/notes="Vector: pT73D-Pac (Pharmacia) with a modified
polylinker; Site_1: EcoRI; Site_2: NotI; sequenced from
forward priming site which reads from the 3' end of the
cDNA. The 4A3A is a directionally cloned and normalized
cDNA library that was constructed from the 4A3A cell line
oligo-T primed cDNA according to: Bonaldo, Lennon & Soares
(1996) : Normalization and Subtraction: Two approaches To
Facilitate Gene Discovery, Genome Research 6, 791-806."
BASE COUNT 167 a 196 c 189 g 146 t 2 others
ORIGIN

Query Match 31.4%; Score 447.8; DB 10; Length 700;
Best Local Similarity 99.1%; Pred. No. 1.3e-115;
Matches 449; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 712 GTGTAGCGGTAGTGGCCACCACCTTCAAGAACTCTGTAGCAGCGCTACATACCTCGCT 771
|||||
Db 599 GTGTAGCGGTAGTGGCCACCACCTTCAAGAACTCTGTAGCAGCGCTACATACCTCGCT 540
|||||
QY 772 CTGCTAATCCTGTTACCAAGTGGCTGCTGCGAGTGGCGATAAGTCGTGTCTTACCGGGTTG 831
|||||
Db 539 CTGCTAATCCTGTTACCAAGTGGCTGCTGCGAGTGGCGATAAGTCGTGTCTTACCGGGTTG 480

```

QY 832 GACTCAAGACGATAGTTACCGGATAAGCCAGCGGTCGGCTGAACGGGGGTTTCGTGC 891
Db |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
479 GACTCAAGACGATAGTTACCGGATAAGCCAGCGGTCGGCTGAACGGGGGTTTCGTGC 420
QY 892 ACACAGCCCAAGCTTGGAGCGAAGACGACTACACCGAAGCTGAGATACCTACAGCGTGAGCAT 951
Db |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
419 ACACAGCCCAAGCTTGGAGCGAAGACGACTACACCGAAGCTGAGATACCTACAGCGTGAGCAT 360
QY 952 TGAGAAAGCCGACGCTTCCGGAAGGGAGAAAGCGGACAGGTATCCGTAAGCGGAGG 1011
Db |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
359 TGAGAAAGCCGACGCTTCCGGAAGGGAGAAAGCGGACAGGTATCCGTAAGCGGAGG 300
QY 1012 GTCGGAACAGGAGAGCGGACGAGGAGGCTTCCAGGGGAAACGCGTGTATCTTTATAGT 1071
Db |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
299 GTCGGAACAGGAGAGCGGACGAGGAGGCTTCCAGGGGAAACGCGTGTATCTTTATAGT 240
QY 1072 CCTGTGGGTTTCGCCACCTCTGACTTGAGCGTCGATTTTGTGATGCTCGTCAGGGGGG 1131
Db |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
239 CCTGTGGGTTTCGCCACCTCTGACTTGAGCGTCGATTTTGTGATGCTCGTCAGGGGGG 180
QY 1132 CGGAGCCTATGSAAGGAAAGCGGACGAGCGGCGCC 1164
Db |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
179 CGGAGCCTATGSAAGGAAAGCGGACGAGCGGCGCC 147

Search completed: January 17, 2002, 08:54:11
Job time: 7123 sec

DEFINITION Sequence 9 from patent US 5518885.

ACCESSION I21124

VERSION I21124.1 GI:1601478

KEYWORDS Unknown.

SOURCE Unknown.

ORGANISM Unclassified.

REFERENCE 1 (bases 1 to 4530)

AUTHORS Raziuddin, . and Sarkar, F.H.

TITLE ERBB2 promoter binding protein in neoplastic disease

JOURNAL Patent: US 5518885-A 9 21-MAY-1996;

FEATURES Location/Qualifiers

source

1. .4530

/organism="unknown"

BASE COUNT 922 a 1382 c 1346 g 880 t

ORIGIN

Query Match 31.5%; Score 669.6; DB 6; Length 4530;

Best Local Similarity 95.2%; Pred. No. 1.5e-155;

Matches 701; Conservative 0; Mismatches 34; Indels 1; Gaps 1;

QY 228 ACCGCAGCAGATGATCATGTCAGACCTGCCCCGGCGCTGGGGCATGGTCCACCA 287

Db 3225 ACCCGACAGGGCTTCTTGTCCAGACCTGCCCGGGCGCTGGGGCATGGTCCACCA 3284

QY 288 CAGGCACCGAGCTCATCTACAGGAGTGCGGTGGGGACCTGACACTAGGGCTGGAGCC 347

Db 3285 CAGGCACCGAGCTCATCTACAGGAGTGCGGTGGGGACCTGACACTAGGGCTGGAGCC 3344

QY 348 CTCTGAAGAGGAGGCCCCCAGGTCTCCTACCTGGCACCTCCGAAGGGGCTGGTCCGATGT 407

Db 3345 CTCTGAAGAGGAGGCCCCCAGGTCTCCTACCTGGCACCTCCGAAGGGGCTGGTCCGATGT 407

QY 408 ATTTGATGTGACCTGGGAATGGGGCAGCCAGAGGGCTGCAAGGCTCCGCCACACATGA 467

Db 3405 ATTTGATGTGACCTGGGAATGGGGCAGCCAGAGGGCTGCAAGGCTCCGCCACACATGA 467

QY 468 CCCCAGCCCTCTACAGGGTACAGTGAGGAGCCACAGTACCCCTGCCCTCTGAGACTGA 527

Db 3465 CCCCAGCCCTCTACAGGGTACAGTGAGGAGCCACAGTACCCCTGCCCTCTGAGACTGA 527

QY 528 TGCTACGTTGCCCTTGCCTGAGAGGGCCCTCTGCTGCTGCCCGACCTGCTGGTGC 647

Db 3585 TGCTACGTTGCCCTTGCCTGAGAGGGCCCTCTGCTGCTGCCCGACCTGCTGGTGC 647

QY 648 CACTCTGGAAGGCCCCAAGACTCTCTCCCGAGGAGAGATGGGGTCTGAAAGAGCTTTT 707

Db 3645 CACTCTGGAAGGCCCCAAGACTCTCTCCCGAGGAGAGATGGGGTCTGAAAGAGCTTTT 707

QY 708 TGCTTTGGGGTGGCGTGGAGAACCCCGAG - ACTTGACACCCCGAGGAGAGCTGCC 766

Db 3705 TGCTTTGGGGTGGCGTGGAGAACCCCGAG - ACTTGACACCCCGAGGAGAGCTGCC 766

QY 767 TCAGCCCCACCTCTCTCTGCTTCCAGCCAGCCCTTCGACACACCTCTATTACTGGAGCA 826

Db 3765 TCAGCCCCACCTCTCTCTGCTTCCAGCCAGCCCTTCGACACACCTCTATTACTGGAGCA 826

QY 827 GGACCCACAGAGGGGGGCTCCACCCAGACCTTCAAGGGACACCTACGGCAGAGAA 886

Db 3825 GGACCCACAGAGGGGGGCTCCACCCAGACCTTCAAGGGACACCTACGGCAGAGAA 886

QY 887 CCCAGAGTACCTGGGCTCTGGAGCTGCCAGTGTGAAGCCTTAAGGGCCATATGGTCAAGTGG 946

Db 3885 CCCAGAGTACCTGGGCTCTGGAGCTGCCAGTGTGAAGCCTTAAGGGCCATATGGTCAAGTGG 946

QY 947 ATGCTTGAAGGAGG 962

Db 3945 CTGATGTCTCTCAGG 3960

RESULT 5

I59745

LOCUS

DEFINITION

Sequence 9 from patent US 5654406.

ACCESSION I59745

VERSION I59745.1 GI:2478377

KEYWORDS Unknown.

SOURCE Unknown.

ORGANISM Unclassified.

REFERENCE 1 (bases 1 to 4530)

AUTHORS Raziuddin, . and Sarkar, F.Hoque.

TITLE Antibody to ERBB2 promoter binding factor

JOURNAL Patent: US 5654406-A 9 05-AUG-1997;

FEATURES Location/Qualifiers

source

1. .4530

/organism="unknown"

BASE COUNT 922 a 1382 c 1346 g 880 t

ORIGIN

Query Match 31.5%; Score 669.6; DB 6; Length 4530;

Best Local Similarity 95.2%; Pred. No. 1.5e-155;

Matches 701; Conservative 0; Mismatches 34; Indels 1; Gaps 1;

QY 228 ACGGCAGCAGAGATCACAATGTCAGACCTGCCCCGGCGCTGGGGCATGGTCCACCA 287

Db 3225 ACCCGACAGGGCTTCTTGTCCAGACCTGCCCGGGCGCTGGGGCATGGTCCACCA 3284

QY 288 CAGGCACCGAGCTCATCTACAGGAGTGCGGTGGGGACCTGACACTAGGGCTGGAGCC 347

Db 3285 CAGGCACCGAGCTCATCTACAGGAGTGCGGTGGGGACCTGACACTAGGGCTGGAGCC 3344

QY 348 CTCTGAAGAGGAGGCCCCCAGGTCTCCTACCTGGCACCTCCGAAGGGGCTGGTCCGATGT 407

Db 3345 CTCTGAAGAGGAGGCCCCCAGGTCTCCTACCTGGCACCTCCGAAGGGGCTGGTCCGATGT 407

QY 408 ATTTGATGTGACCTGGGAATGGGGCAGCCAGAGGGCTGCAAGGCTCCGCCACACATGA 467

Db 3405 ATTTGATGTGACCTGGGAATGGGGCAGCCAGAGGGCTGCAAGGCTCCGCCACACATGA 467

QY 468 CCCCAGCCCTCTACAGCGGTACAGTGAGGAGCCACAGTACCCCTGCCCTCTGAGACTGA 527

Db 3465 CCCCAGCCCTCTACAGCGGTACAGTGAGGAGCCACAGTACCCCTGCCCTCTGAGACTGA 527

QY 528 TGCGTACGTTGCCCTTGCCTGAGAGGGCCCTCTGCTGCTGCCCGACCTGCTGGTGC 587

Db 3525 TGCGTACGTTGCCCTTGCCTGAGAGGGCCCTCTGCTGCTGCCCGACCTGCTGGTGC 587

QY 588 TGCGCCCCAGCCCTTGCCTGAGAGGGCCCTCTGCTGCTGCCCGACCTGCTGGTGC 647

Db 3585 TGCGCCCCAGCCCTTGCCTGAGAGGGCCCTCTGCTGCTGCCCGACCTGCTGGTGC 647

QY 648 CACTCTGGAAGGCCCCAAGACTCTCTCCCGAGGAGAGATGGGGTCTGAAAGAGCTTTT 707

Db 3645 CACTCTGGAAGGCCCCAAGACTCTCTCCCGAGGAGAGATGGGGTCTGAAAGAGCTTTT 707

QY 708 TGCGTTTGGGGTGGCGTGGAGAACCCCGAG - ACTTGACACCCCGAGGAGAGCTGCC 766

Db 3705 TGCGTTTGGGGTGGCGTGGAGAACCCCGAG - ACTTGACACCCCGAGGAGAGCTGCC 766

QY 767 TCAGCCCCACCTCTCTCTGCTTCCAGCCAGCCCTTCGACACACCTCTATTACTGGAGCA 826

Db 3765 TCAGCCCCACCTCTCTCTGCTTCCAGCCAGCCCTTCGACACACCTCTATTACTGGAGCA 826

QY 827 GGACCCACAGAGGGGGGCTCCACCCAGACCTTCAAGGGACACCTACGGCAGAGAA 886

Db 3825 GGACCCACAGAGGGGGGCTCCACCCAGACCTTCAAGGGACACCTACGGCAGAGAA 886

QY 887 CCCAGAGTACCTGGGCTCTGGAGCTGCCAGTGTGAAGCCTTAAGGGCCATATGGTCAAGTGG 946

Db 3945 CTGATGTCTCTCAGG 3960

Db 3885 CCCAGAGTACCTGGGTCTGGACGTGCTCCAGTGTGAACCAAGAGGCCAAGTCCGAGAAGCC 3944

QY 947 ATGCGCTTGACCCAGG 962

Db 3945 CTGATCTGTCTCAGG 3960

RESULT 6

HUMHER2A 4530 bp mRNA PRI 18-SEP-1995

LOCUS Human tyrosine kinase-type receptor (HER2) mRNA, complete cds.

DEFINITION M11730

ACCESSION M11730.1 GI:183986

VERSION tyrosine kinase.

KEYWORDS Homo sapiens (clone: lambda-HER2-436) fetal cDNA to mRNA.

SOURCE Homo sapiens

ORGANISM Homo sapiens

REFERENCE 1 (bases 1 to 4530)

AUTHORS Coussens, L., Yang-Feng, T.L., Liao, Y.-C., Chen, E., Gray, A., McGrath, J., Seeburg, P.H., Libermann, T.A., Schlessinger, J., Francke, U., Levinson, A. and Ullrich, A.

TITLE Tyrosine kinase receptor with extensive homology to EGF receptor shares chromosomal location with neu oncogene

JOURNAL Science 230 (4730), 1132-1139 (1985)

MEDLINE 86070181

REFERENCE 2 (bases 1701 to 1719)

AUTHORS Ullrich, A.

JOURNAL Unpublished (1988)

FEATURES

Location/Qualifiers

1..4530

/organism="Homo sapiens"

/db_xref="taxon:9606"

/clone="lambda-HER2-436"

/dev_stage="fetal"

<1..4530

/note="HER2 mRNA"

151..3918

/note="HER2 receptor"

/codon_start=1

/protein_id="AA75493.1"

/db_xref="gi:306840"

/translation="METAALCRWGLLALLPPLPAASTOVCTGDMKRLPASPETHLD MLRLHYQCVQVQGNLELTLYLTPASFLQDIOEVQGYVLIAHNRQVPLQRLRIV RGTOLFEDNYALVDLNDGDLNNTPTVTSASGGLRELQLRLSLTEILKAGVLIQNPQ LCYQDTLLKWDIFHNKQALTLIDTNRSRACPCSPMKCSRWGSESEDCQSLTRYT VCAGCARCKPLPTDCEQCACTGPKRHSDDLCLAHFNHSGICEHCHCALVTYNT DTFESMPNPEGRYTFGASCTVACPYNYLSTDVSGCTVCLPHNQEVTAEDGTQREKC SKPCARVYGLGMEHLREVRATVSNIOEFAGCKIFGLAFLEPSFDGDPASNTAPL QPEQLQVPTLEEITGYLIISAWPDLPLSVFNQLOVIRGRIHLHNGAYSLTLOGLI SWGLRLSRELSGLALHNNHLCFVHTVPWDLFRNPHOALLHTANRPEDECVGEG LACHQLCARGHCWGPGTQCVCNQSLFQGECEVRCVLOGLPREYVNAHCLCPHE CQPQNGSVTCFEPADQCACAHYKDPFCVAFCSGVKPKDLSYMPINWFDEEGACQ PCPINCTHSCVDLDDKGCAPORASPLTFSIVSAVVILLVVLVGVFGLILKRRQKI RKYTMRLRLQETLELTPSGAMPNQAMRIKLETRELKVKVLSGAGFYVYGIWI PDGENKIPVAIKVLENTSPKANKEILDEAYVMAGVSPYVSRLLGLCLSTVOLVT QLMVPGCLLDHVRNRLSGDILLNWCQIAKMSYLEDVRLVHRDLAARNLVKSP NHVKTQDGLARLLIDFTEYHADGKVPYIKWMALESILRRRTHQSDWNSGYTWLE LMTGAKPYDGPAREIPDLLEKGERLPQPICTIDVTYMWKMDISECRPREL VSEFSRMARDPQRFVIONEDLGPASPLDSTFYSLDEDDMDLDVAEEYLVPOQGF FCPDPAPAGGMVHRHRSSTRSGGDLTLGLEPSEEAAPRSLAPSEAGSDVFDG DLGMAAGLQSLPTSPQRYSEDPVLPSTDGYVAPLTCSPQPEYVNPQDVR POPPSRGGPLPAARPAATLERAKTISPGKNGVVKDVFAGGAVENPEYLTPOGGAA POPHPPPAFSFDNLYWDQDPPPERGAPSTFKGTPTAENPEYLGLDVVPV"

old_sequence 1701..1719

/citation=[1]

BASE COUNT 922 a 1382 c 1346 g 880 t

ORIGIN Chromosome 17q21-q22.

Query Match 31.5%; Score 669.6; DB 9; Length 4530;

Best Local Similarity 95.2%; Pred. No. 1.5e-155;

Matches 701; Conservative 0; Mismatches 34; Indels 1; Gaps 1;

QY 228 ACGGACGAGAGAATCATCATGTCCAGACCCCTCCCGGGCGCTGGGGGATGTTCCACCA 287

Db 3225 ACCCAGCAGAGGGCTTCTTCTGTCTCCAGACCCCTCCCGGGCGCTGGGGGATGTTCCACCA 3284

QY 288 CAGGACCCGACGCTCATCTACAGAGGTGGGGTGGGGACCTGACACTAGGGGTGGAGCC 347

Db 3285 CAGGACCCGACGCTCATCTACAGAGGTGGGGTGGGGACCTGACACTAGGGGTGGAGCC 3344

QY 348 CTTCTGAAGAGGAGGCCCCCAGAGTCTCCACTGGCACCTCCGAAGGGCTGGCTCCGATCT 407

Db 3345 CTTCTGAAGAGGAGGCCCCCAGAGTCTCCACTGGCACCTCCGAAGGGCTGGCTCCGATCT 3404

QY 408 ATTGTATGTGACCTGGGAATGGGGACGCAAGGGGCTGCAAGGCTCCCCACACATGA 467

Db 3405 ATTGTATGTGACCTGGGAATGGGGACGCAAGGGGCTGCAAGGCTCCCCACACATGA 3464

QY 468 CCCAGCCCTCTACAGCGGTACAGTAGGAGGCCACAGTACCCCTGCCCTCTGTAGACTGA 527

Db 3465 CCCAGCCCTCTACAGCGGTACAGTAGGAGGCCACAGTACCCCTGCCCTCTGTAGACTGA 3524

QY 528 TGGCTACGTTGCCCCCTGACCTGCAGCCCCCAGCGCTGAATGTGAACGACGACGATGT 587

Db 3525 TGGCTACGTTGCCCCCTGACCTGCAGCCCCCAGCGCTGAATGTGAACGACGACGATGT 3584

QY 588 TCGGCCCCAGCCCCCTTCGCCCGCAGAGGGCCCTCTGCCTGTGCCCGACCTGCTGTTGC 647

Db 3585 TCGGCCCCAGCCCCCTTCGCCCGCAGAGGGCCCTCTGCCTGTGCCCGACCTGCTGTTGC 3644

QY 648 CACTCTGGAAGGCCCAAGACTCTCTCCCAGGGAAGATGGGTCTGTCACAAAGACGTTTT 707

Db 3645 CACTCTGGAAGGCCCAAGACTCTCTCCCAGGGAAGATGGGTCTGTCACAAAGACGTTTT 3704

QY 708 TGCCTTTGGGGTGGCGTGGAGAACCCCGAG-ACITGACACCCCGAGGAGGAGCTGCCCC 766

Db 3705 TGCCTTTGGGGTGGCGTGGAGAACCCCGAGTACTTGACACCCCGAGGAGGAGCTGCCCC 3764

QY 767 TCAGCCCCACCTCTCTCTCTTCAGCCCGACGCTTCGACAACTCTATTACTTGGGACCA 826

Db 3765 TCAGCCCCACCTCTCTCTCTTCAGCCCGACGCTTCGACAACTCTATTACTTGGGACCA 3824

QY 827 GGACCCACGAGCGGGGGCTCCACCCAGCACCTTCAAGAGGACACCTACGCGCAGAGAA 886

Db 3825 GGACCCACGAGCGGGGGCTCCACCCAGCACCTTCAAGAGGACACCTACGCGCAGAGAA 3884

QY 887 CCCAGAGTACCTGGGTCTGGAGTGGCCAGTGTGAAGCCCTTAAGGGCCCATATGGTGAGTGG 946

Db 3885 CCCAGAGTACCTGGGTCTGGAGTGGCCAGTGTGAAGCCCTTAAGGGCCCATATGGTGAGTGG 3944

QY 947 ATGCTTTGACCCCGAGG 962

Db 3945 CTGATGTGTCTCAGG 3960

RESULT 7

AR034479

LOCUS AR034479 3768 bp DNA PAT 29-SEP-1999

DEFINITION Sequence 1 from patent US 5869445.

ACCESSION AR034479

VERSION AR034479.1 GI:5950084

KEYWORDS

SOURCE Unknown.

ORGANISM Unknown.

REFERENCE 1 (bases 1 to 3768)

AUTHORS Cheever, M.A. and Disis, M.L.

TITLE Methods for eliciting or enhancing reactivity to HER-2/neu protein

JOURNAL Patent: US 5869445-A 1 09-FEB-1999;

FEATURES

Location/Qualifiers

1..3768

/organism="unknown"

BASE COUNT 759 a 1171 c 1119 g 719 t

ORIGIN

[illegible]

Direct Submission
Submitted (23-OCT-1997) to the DDBJ/EMBL/GenBank databases. Hiroshi Yokota, Rakuno Gakuen University, Veterinary Biochemistry, Bunkyo-daimeidomori, Ebetsu, Hokkaido 069, Japan
(E-mail: BXA03502@niftyserve.or.jp, tel: 011-386-1111, Fax: 011-387-5890)
2 (bases 1 to 3780)
Yokota, H.
cDNA cloning of erbB-2 from canine mammary gland
Published Only in Database (1997) In press
Location/Qualifiers
1. 3780
/organism="Canis familiaris"
/db_xref="taxon:9615"
1. 3780
/codon_start=1
/product="erbB-2"
/protein_id="BAA33127.1"
/db_xref="GI:2575867"
/translation="MELAAWCRWGLLLALLPSGAAGTQVCTGDMKRLRLPASPTHLD MLRLHYQCVQVVOGNLELTVLPANASLFLQIOEVGYVLIASHQVROIPLQRLRV RGTOLFEDNVALVLDNGDPLEGGIPAPGAQGLRELQRLSLTEILKGGVLIQSRPQ LCHQDTLWQDVFHKNQALTLIDTRNFSNCPSPACKDAHGWAGSSGBCQSLRT VCAGGACRGQPQTDCCHEQCAAGCTYKPSHSDCLACHFNHSGICEICELCPALTYNT DTFESNPERYTFGASCVTSCPNYLDVSGTCLVCLPLNQEVTAEDGTQCEK SKPCARVCYGLMEHLREAVRANTSANIQEFAGCKKIFGSLAFLPESFDGDPASNTAPL OPEOLRVFEALGETTYLXISAWPDSLNLVSFONLVRIGRVLHDGAYSLTLOGLI SWGLRSRLRGSLALHNRNRLCFVHTVPMDQLFRNPHQALLHSANRPREECVGG LACTPCAHCWGPGPTQCVNSQFLRQGECEVRCVLQGLPRVYIKDCLPCHSEC QPQNGVTCSEADQVACAHYKDPFCVARCPGSKVPDLSEMPDKWFADEETCOQ CPINTHSCADLDKGPAPQORASPVTSIIAAVVGILLVAVVGLVGLILIKRRQKIR KYTMRLLQETELVEPLTPSGAMPNQMRILKTELKRVKLVGAGFTVYKGIWIP DMGVKIPVAIKVURENTSPKANKEILDEAVMAGVSPVYSLILGLICTSTVOLVTC LGPNVCLLDHVRHGRBLGSQDLLNWCQVIAKMSYLEDVRLVLDLAARNVLYKSPN HVKITDGLARLLDIDETEHADGKVPKIMMALESIPPRFTHQSDVMSTGVTWEL MTFGAKYDGPAREIPDLLEKGERLPOPICTIDVIMYIMVKWMDSECRPRELY AEFSRMADQRFVIEIDLEGPASPLDSTFYRLSLEDMDMDGLVDAEYLVLPQQGF CPEPTGAGGTASHRRSSSTRNGSGELTGLESEEPKPSPLAPSEAGSDVDFGD LGMAAGLQSLQSDPSPLQRYSEDVTPVLPPLPDKVAPLTCSPQEVVNOPEVMP OPPLALRGLPLPPSPACATLERPKTSLPKTLSPKNGVVKVDFAFGSAVPEYLAPR GRAAPQHPHPAPSPAFDNLVYWDQDSERGSPPSTEGPTTAENPEYLGUDVPV"

BASE COUNT 743 a 1162 c 1150 g 725 t
ORIGIN

Query Match 24.3%; Score 516.2; DB 4; Length 3780;
Best Local Similarity 85.3%; Pred. No. 1.8e-117;
Matches 605; Conservative 0; Mismatches 88; Indels 16; Gaps 2;

QY 228 ACGGCAGCAGAGATGCATGTCCAGACCCCTGCCCGGCGCTGGGGGCGATGGTCCACCA 287
DB 3072 ACCCCAGCAGGGTTCTTCTGCCAGAACCTACCCAGGGGCTGGGGCAGCTGCCACCG 3131

QY 288 CAGGCACCGCAGCTCATCTACACGAGGTGGGGTGGGACCTGCACATGAGGCTGGAGCC 347
DB 3132 ACGGCACCGCAGCTCATCCAGCAATGGGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 3191

QY 348 CTCTCAAGAGGAGGCCCCCAGCTCTCCACTGCGACCCCTCCGAGGGGCTGGCTCGATGT 407
DB 3192 CTCCAGGAGGAGGAGCCCCCAAGTCTCCACTGGACCCCTCAGAGGGGCTGGCTGGCTG 3251

QY 408 ATTTGATGTGACCTTGGGAATGGGGGAGCCCAAGGGGCTGCAAGAGGCTCCCCACACATGA 467
DB 3252 GTTTCATGTGACTTGGGAATGGGGGAGCCCAAGGGGCTGCAGAGCTTCCCTCAGAGA 3311

QY 468 CCCCAGCCCTTACAGCGGTACAGTGAGAGACCCCAAGTACCCCTGCCCTCTGAGACTGA 527
DB 3312 CCCCAGCCCTTCCACGCGGTACAGTGAGAGACCCCTACCGTACCCCTTGGCCCTTGAGACTGA 3371

QY 528 TGGCTACCTTGGCCCTTACCTGACCTGACCTGACCTGACCTGACCTGACCTGACCTGAC 587
DB 3372 TGGTAAGGTTGCCCCCTTACCTGACCTGACCTGACCTGACCTGACCTGACCTGACCTG 3431

QY 588 TCGGCCCCAGCCCTTGGCCCGCAGAGGGCCCTCTGCTGCTGCCCGACCTGCTGGTGC 647
DB 3432 TTGGCCGCGAGCCCTTGGCCCTAGAGGGCCCTTGGCTCTTCCCGACCGCTGGTGC 3491

QY 648 CACTCTGGAAAGG-----CCCAAGACTCTCTCTCCCGAGGAAGAATGGGT 692
DB 3492 CACTCTGGAAAGGCCCAAGACTCTCTCTCCCGAAGACTCTCTCTCCCGAAGAATGGGT 3551

QY 693 CQTCAAGACGTTTTTGGCTTTGGGGGTGCGCTGGAGAACCCCGAG-ACCTTACACCCCA 751
DB 3552 TGTCAAGACGTTTTTGGCTTTGGGAGTGTCTGTGGAGAAATCCGGAGTACTTGGACCCG 3611

QY 752 GGGAGGAGTGCCTTCAGCCCGCCACCTCTCTCTGCTTTCAGCCCGACCTTTCAGAACCT 811
DB 3612 GGGAGAGCTGCCCCCTCAGCCCCACCTCTCTCTCAGCCCTTTCAGCCCGACCTTTCAGAACCT 3671

QY 812 CATTACTGGGACAGGACCCAGAGAGGGGGGGCTCCACCCAGACCTTCAAGAGGAC 871
DB 3672 GTATTACTGGGACAGGATCCATCAGAGCGGGGCTCTCCACCCAGACCTTTCAGAGGAC 3731

QY 872 ACCTACGCGAGAACCCAGAGTACTTGGGTCTGGACGTGCCAGTGTGA 920
DB 3732 CCTACAGAGAAACCCGGAGTACTTGGGGCTGGACGTGCCAGTGTGA 3780

RESULT 10
AX189662 LOCUS AX189662 3771 bp DNA PAT 08-AUG-2001
DEFINITION Sequence 1 from Patent WO0148205.
ACCESSION AX189662
VERSION AX189662.1 GI:15143051
KEYWORDS house mouse.
SOURCE Mus musculus
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE 1 (bases 1 to 3771)
AUTHORS Spies, A.G.
TITLE Murine neu sequences and methods of use therefor
JOURNAL Patent: WO 0148205-A 1 05-JUL-2001;
CORIXA CORPORATION (US)
FEATURES
Location/Qualifiers
1. 3771
/organism="Mus musculus"
/db_xref="taxon:10090"
BASE COUNT 808 a 1105 c 1068 g 790 t
ORIGIN

Query Match 23.9%; Score 508.8; DB 6; Length 3771;
Best Local Similarity 82.3%; Pred. No. 1.2e-115;
Matches 596; Conservative 0; Mismatches 127; Indels 1; Gaps 1;

QY 198 GGGGTGGTCTTTTGGATCTCTCATCAAGCGACGCGGAGAGATCACATGTCCAGACC 257
DB 3048 GGAGTGTGATGCTGAAGAGTACTTGTGTACCCCGCAGGAGTCTTCTCCCGACCC 3107

QY 258 TGCCCGGGCGCTGGGGGCGATGGTCCACACAGGACCCCGACCTATCTACAGAGTGG 317
DB 3108 TGCCCTAGGTACTTGGGAGCACAGCCCGCAGACCCCGACCTCGTCGGCCAGGAGTGG 3167

QY 318 CGTGGGACCTGACACTAGGCGCTGGAGCCCTTGAAGAGGAGGCCCGCAGGTCTCCACT 377
DB 3168 CGTGGTGAAGTGAAGTGGGCGCTGGAGCCCTTGAAGAGGAGGCCCGCAGGTCTCCACT 3227

QY 378 GGCACCCCTCCGAAGGGCTGGCTCCGATGATTTGATGGTACCTGGGAATGGGGGACG 437
DB 3228 GGCTCCCTCCGAAGGGCTGGCTCCGATGATTTGATGGTACCTGGGAGTGGGGTAAC 3287

QY 438 CAAAGGGCTGCAAGGCTCCCGACACATGACCCCGCCTTACAGCGGTACAGTGA 497
DB 3288 CAAAGGACTGACAGGCTCTCTCCACATGACCTCAGCCCTTACAGCGGTACAGTGA 3347

Qy	498	CCCACAGTACCCCTGCCCTCTGAGACTGATGGCTACGTTGCCCCCTCTGACCTGCAGGCC	557
Db	3348	TCCACACATTACCTCTGCCCCCGAGACTGATGGCTACGTTGCTCCCTTGGCTTCAGGCC	3407
Qy	558	CCAGCCTGAATATGTGAACAGCAGCATGTCGGCCCGACGCCCTTCGCCCCGAGAGGG	617
Db	3408	CCAGCCGAGTATGTGACACGACGAGGTCGGCCTCAGTCTCCCTTTCAGCCCGAGAGGG	3467
Qy	618	CCCTCTGCCTGCTGCCGACCTGCTGGTGCCACTCTGAAAGGCCAAGACTCTCTCCCC	677
Db	3468	TCCCTCCGCTCCCATCCGACCTGCTGGTGCTACTCTAGAAAGACCCAGACTCTCTCTCC	3527
Qy	678	AGGGAAGAATGGGGTCGTCAAGACGTTTTTGCCTTTGGGGTGCCGTGGAGAACCCCG	736
Db	3528	TGGGAAAAATGGGGTTGTCAAGACGTTTTTGCCTTTGGGGTGCTGTGGAGAACCCCTGA	3587
Qy	737	AGACTTGACACCCACGAGGAGCTGCCCTCAGGCCACACCTCCTCCTGCTTCAGGCC	796
Db	3588	ATACCTAGCACCCAGAGCAGGCACTGCCTCTCAGCCCCACCCCTTCTCTGCTTCAGGCC	3647
Qy	797	AGCCTTCGACAACCTCTATTACTTGGACAGGACCCACAGAGGGGGGGTCCACCCAG	856
Db	3648	AGCCTTTGACAACTCTATTACTTGGACACAGAACTCATCGGAGCAGGGTCTCTCCACCAAG	3707
Qy	857	CACCTTCAAAAGGACACCTACGCGCAGAGAACCCAGGTACCTGGTCTGTGGACGTGCCAGT	916
Db	3708	TACCTTTGAAGGACCCCCACTGCAGAGAACCTCAGTACCTAGGCCCTGGATGTGCCAGT	3767
Qy	917	GTGA 920	
Db	3768	ATGA 3771	

RESULT	11
I21129	
LOCUS	3955 bp DNA
DEFINITION	Sequence 14 from patent US 5518885.
ACCESSION	I21129
VERSION	I21129.1 GI:1601483
KEYWORDS	
SOURCE	Unknown.
ORGANISM	Unknown.
REFERENCE	Unclassified.
AUTHORS	1 (bases 1 to 3955)
TITLE	Razidudin;; and Sarkar,F.H.
JOURNAL	ERBB2 promoter binding protein in neoplastic disease
FEATURES	Patent: US 5518885-A 14 21-MAY-1996;
	Location/Qualifiers
	1..3955
	/organism="unknown"
BASE COUNT	842 a 1147 c 1136 g 830 t
ORIGIN	
Query Match	23.3%; Score 495.8; DB 6; Length 3955;
Best Local Similarity	83.0%; Pred. No. 2.1e-112;
Matches 577; Conservative	0; Mismatches 117; Indels 1; Gaps 1;

RESULT	12
LOCUS	I59750
DEFINITION	159750 bp DNA PAT 07-OCT-1997
ACCESSION	Sequence 14 from patent US 5654406.
VERSION	I59750.1 GI:2478382
KEYWORDS	.
SOURCE	Unknown.
ORGANISM	Unknown.
FEATURES	Unclassified.
AUTHORS	1 (bases 1 to 3955)
TITLE	Raziuddin;;. and Sarkar,F.Hoque.
JOURNAL	Antibody to ERBB2 promoter binding factor
Patent:	US 5654406-A 14 05-AUG-1997;
Location/Qualifiers	
1..3955	
/organism="unknown"	
BASE COUNT	842 a 1147 c 1136 g 830 t
ORIGIN	
Query Match	23.3%; Score 495.8; DB 6; Length 3955;
Best Local Similarity	83.0%; Pred. No. 2.1e-112;
Matches 577; Conservative	0; Mismatches 117; Indels 1; Gaps
QY 232	CAGCAGAAGATCACATGTCCAGACGCTGCCCGGGCGCTGGGGCATGTGTCACACAGG 291
Db 3110	CAGCAGGATTCTTCTCCC CGACCTACCC CAGGC ACTGGGAGCAGAGCCCATAGAAGG 3169
QY 292	CACCGCAGCTCACTACACAGAGTAGTGCGGTGGGGACTGTACACTAGGCGCTGGAGCCCTCT 351
Db 3170	CACCGCAGCTCGTCTCCACCAGAGTAGTGAGTTGGTGAAGTGTGACACTGGGCTGGAGCCCTCG 3229
QY 352	GAGAGGAGGCCCCCAGGTCCTCACTGGCACCCCTCCGAGGGGCTGGCTCCGATGTATTT 411
Db 3230	GAGAAGGGCCCCCAGATCTCCA CTGGCTCCCTCGGAAGGGGTGGCTCCGATGTGTTT 3289
QY 412	GATGGTGACCTGGGAATGGGGCAGCCAAAGGGGCTGTCAAAGCCCTCCCAACACATGACCCC 471


```
QY 532 TAGTTGGCCCCCTGACCTGCAGCCCGAGCGTGAATATGTGAACCCAGACATGTTGG 591
Db 3410 TATGTTGCTCCCTGGCTGCAGCCCGAGCGTGAATATGTGAACCCAGACATGTTGG 3469
QY 592 CCCCAGCCCCCTTCCCGCGAGAGGGCCCTGTGCTGCTGCCCGACCTGTGTCGCCACT 651
Db 3470 CTTAGCCCTTCTTAACCCAGAGGGTCTCTGCTGCTGCTGCCCGACCTGTGTCGCCACT 3529
QY 652 CTGGAAGGCCCAAGACTCTCTCCCGAGGAAGTGGGTCTGCAAGAGGTTTGGCC 711
Db 3530 CTAGAAGACCAAGACTCTCTCTCTGGGAAGTGGGTCTGCAAGAGGTTTGGCC 3589
QY 712 TTTGGGGTGGTGGGAACCCCGC-AGACTGTACACCCCGAGGAGGAGCTGCCCTCAG 770
Db 3590 TTCGGGGTGGTGGGAACCCCTGAACTACTTAGTACCGAGAGGACCTGCTCTCCG 3649
QY 771 CCCACCCCTCTCTGCTTTCAGCCCGAGCTTCGACACCTCTATTACTGGGGACAGGAC 830
Db 3650 CCCACCCCTCTCTGCTTTCAGCCCGAGCTTCGACACCTCTATTACTGGGGACAGGAC 3709
QY 831 CCACAGAGGGGGGCTCCACCCAGACCTTCAAAGGACACCTACGGCAGAGAACCCA 890
Db 3710 TCATCGAGGAGGGGGCTCCACCAAGTAACTTTGAAGGGACCCCACTGCAGAGAACCT 3769
QY 891 GAGTACCTGGGTCTGGAGCGCCAGTGTGAAGCCT 925
Db 3770 GAGTACCTGGGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 3804

RESULT 14
LOCUS HAMNEU 4062 bp mRNA ROD 01-FEB-2000
DEFINITION Syrian golden hamster neu mRNA for p-185, complete cds.
ACCESSION D16295
VERSION D16295.1 GI:493236
KEYWORDS EGF-receptor; cellular oncogene; neu-differentiation
factor/herregulin; p-185; tyrosine kinase.
SOURCE Mesocricetus auratus (individual isolate animal 14) peripheral
nerve neoplastic Schwann cell cell-line 14-2 (library: NIH3T3) cDNA
to mRNA, clone pSHN.
ORGANISM Mesocricetus auratus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Cricetinae;
Mesocricetus.
REFERENCE 1 (bases 1 to 4062)
Nakamura,T., Ushijima,T., Ishizaka,Y., Nagao,M., Arai,M.,
Yamazaki,Y. and Ishikawa,T.
Cloning and activation of the Syrian hamster neu proto-oncogene
Gene 140 (2), 251-255 (1994)
94193007
2 (bases 1 to 4062)
Nakamura,T.
Direct Submission
Submitted (19-MAY-1993) to the DDBJ/EMBL/GenBank databases. Takuro
Nakamura, Faculty of Medicine, University of Tokyo, Department of
Pathology, 7-3-1 Hongo, Bunkyo-ku, Tokyo, 113, Japan
(Tel:03-3812-2111(ex.3356), Fax:03-3815-8379)
Submitted (19-May-1993) to DDBJ by:
Takuro Nakamura
Department of Pathology
Faculty of Medicine, University of Tokyo
7-3-1 Hongo
Bunkyo-ku Tokyo 113
Japan
Phone: 03-3812-2111 x3356
Fax: 03-3815-8379.
Location/Qualifiers
1. .4062
/organism="Mesocricetus auratus"
/isolate="animal 14"
/db_xref="taxon:10036"
/cell_line="14-2"

FEATURES
source
```

```
/cell_type="neoplastic Schwann cell"
/clone_lib="NIH3T3"
/tissue_type="peripheral nerve"
```

```
33..3797
/gene="neu"
sig_peptide
```

```
33..86
/gene="neu"
/product="putative"
CDS
33..3797
/gene="neu"
/codon_start=1
```

```
/product="p-185 precursor"
/protein_id="BAA03801.1"
/db_xref="GI:747595"
```

```
/translation="MELAAWCGMGLLALLSPCAGTQVCTGDMKRLPASPETHLD
IVRLHVGCGVOVGNLELYLPANATLSFLQDIOEVQGLMLIAHSHVRHVPQRLRIY
RGYQLPDKYALAVLDRPDLNVTTATGTPEGLRELQRLSLTELKGLVLRNPQ
LCYQDYLWKDVRKNNQLAPVDIDNRSRACPPCAPKDNHNGASPEDCQTLTGT
IAPRAAARLPTDCHEQCAAGCTGPKHSDCLACLFHNSGICELHCPALVTNT
DTPESPNEGRYTFGASCTTCCPYNLSTEVGCTLVCLNNQVTAEDGTORCEK
TPQLQVFTLEITGLYISAWPDSLHLSVFQNLVTRGVLDHSGAYSLAQGLGI
RWLGLSLRELGLVLIHRNTHLCFVHTVPMDQLFRNPHQALLHSGNPSEEGCKLD
FACYPICAHGCHGPGPQCVCNCSHFLGQCCKECRVWKGLPREYVNGKCLCPHE
COPQNSTECTSEADQCTCAHYKDSPPFCVACRSPKPDLSYMPIWKYPDERMCQ
PCDINCTSCVDLDERGCPAEORASPATSIATVVVGIILFLVGVVGLIKRRKQI
RAYTMRLLQETELVEPLTSGAMPNQAMRILKTELKRVKVLGSGAGFVYKGIWI
PDGENKIPVAIKVLENTSPKANKEILDEAYVMAGISGVPYVSLRGLCTLTSTVQLT
OLMPYGLDHLVREHGRIGSODLLNWCVOAKGMSYLEYVLRVHDLAARVLKSP
NHVITDFGLARLDDIETEHADGKVPKIKWIALESILRRVTHSDVNSGTGVVME
LMTFGAPYDGPAREIPDLLEKGRPLPOPICTIDVIMVWKMDIDSECRPREL
VSESRMADPQFVVIONEDLGPSSLDSTFYKSLLEDDMDGLDAEYLVPQGF
FEPDAPGAGSTAHRRHSSTRSGGELTGMSEPPSEPPSPPLAFSEGAGSDVFG
ELMGATGPOSISPRDLSPLQVSEDTPLPTETDGYVAPLACSPQEVVNOEVR
PQPLTPPEGPUPPVPRGATLPERKTLSPKNGVVKDVFVFGGAVENPEYLVPRGSA
SOPHPPLCPAFDNLVYWDQPSRSGSPNTEGTPTAENPEYLGLDVVP"
```

```
mat_peptide
```

```
87..3794
/gene="neu"
/product="p-185"
```

```
misc_feature
```

```
516..1988
/gene="neu"
/note="extracellular domain"
```

```
misc_feature
```

```
1989..2057
/gene="neu"
/note="putative; transmembrane domain"
```

```
misc_feature
```

```
2058..3794
/gene="neu"
/note="cytoplasmic domain"
```

```
BASE COUNT 845 a 1209 c 1182 g 826 t
ORIGIN
```

```
Query Match 22.9%; Score 487; DB 10; Length 4062;
Best Local Similarity 82.4%; Pred. No. 3.2e-110;
Matches 571; Conservative 0; Mismatches 120; Indels 2; Gaps 1;
```

```
QY 232 CAGCAGAGATCATGTCCAGACCCCTGCCCGGGCGCTGGGGCGATGGTCCACACAG 291
```

```
Db 3111 CAGCAGGGAATCTTCTTCCAGACCCCTGCCCGGGCGCTGGGGCGATGGTCCACACAG 3170
```

```
QY 292 CACCCGACCTCATCTACAGAGTGGCGGTGGGACCTGACATAGGCGTGGAGCCCTCT 351
```

```
Db 3171 CACCGTAGCTCGTCAACAGAGTGGAGCGGCGGAGCTGACATGGGCGATGGAGCCCTCG 3230
```

```
QY 352 GAACAGAGGCCCCCAGGTCTCCACTGGCACCCCTCCGAGGGGCTGGCTCCGATGATTT 411
```

```
Db 3231 GGAGAGAGCCCCCAGGTCTCCACTGGCTCCCTCGAAGGGGCTGGCTGATGTGTTT 3290
```

```
QY 412 GATGCTGACCTGGGAATGGGCGAGCAGGCGCTGCAAGCCCTCCCCACAGATGACCCC 471
```

```
Db 3291 GAGGTGAACCTGGGATGGGCGCTACCAAGGGCGCGAGCATCTCTCCACGTGACCTC 3350
```

```
QY 472 AGCCCTCTACAGCGGTACAGTGAAGGACCCCAAGTACCCCTGCGCTCTGAGACTGATGC 531
```

Qy	1591	ACACAGCCAGCTTGGAGCGAAGCAGCCTACACCGAACTGAGATACCTACACGGTGAGCAT	1650
Db	806	ACACAGCCAGCTTGGAGCGAAGCAGCCTACACCGAACTGAGATACCTACACGGTGAGCAT	747
Qy	1651	TGAGAAAGCGCCACGCTTCCGGAAGGGAGAAAGCGGACAGGTATCCGGTAAAGCGGCAGG	1710
Db	746	TGAGAAAGCGCCACGCTTCCGGAAGGGAGAAAGCGGACAGGTATCCGGTAAAGCGGCAGG	687
Qy	1711	GTCCGAACAGGAGAGCGCAGAGGAGCTTCCAGGGGGAAACGCTGTATCTTTATAGT	1770
Db	686	GTCCGAACAGGAGAGCGCAGAGGAGCTTCCAGGGGGAAACGCTGTATCTTTATAGT	627
Qy	1771	CTGTCTGGGTTTCGCCACACTCTGACTTGAGCGTCGATTTTGTGATGCTCGTCAGGGGGG	1830
Db	626	CTGTCTGGGTTTCGCCACACTCTGACTTGAGCGTCGATTTTGTGATGCTCGTCAGGGGGG	567
Qy	1831	CGGAGCCTATGGAACACGCCACGACGCGCGCGGGG	1868
Db	566	CGGAGCCTATGGAACACGCCACGACGCGCGCGGGG	529

Search completed: January 17, 2002, 11:49:53
Job time: 17205 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: January 17, 2002, 12:03:06 ; Search time 599.86 Seconds
(without alignments)
3037.069 Million cell updates/sec

Title: US-09-242-202A-22

Perfect score: 2125

Sequence: 1 GCCACCATGCCCTGACCT.....AGCCTCTCCACAGGTACCT 2125

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 930621 seqs, 428662619 residues

Total number of hits satisfying chosen parameters: 1861242

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : N_Geneseq_1101.*
1: /SID52/gcgdata/geneseq/geneseq/NA1980.DAT.*
2: /SID52/gcgdata/geneseq/geneseq/NA1981.DAT.*
3: /SID52/gcgdata/geneseq/geneseq/NA1982.DAT.*
4: /SID52/gcgdata/geneseq/geneseq/NA1983.DAT.*
5: /SID52/gcgdata/geneseq/geneseq/NA1984.DAT.*
6: /SID52/gcgdata/geneseq/geneseq/NA1985.DAT.*
7: /SID52/gcgdata/geneseq/geneseq/NA1986.DAT.*
8: /SID52/gcgdata/geneseq/geneseq/NA1987.DAT.*
9: /SID52/gcgdata/geneseq/geneseq/NA1988.DAT.*
10: /SID52/gcgdata/geneseq/geneseq/NA1989.DAT.*
11: /SID52/gcgdata/geneseq/geneseq/NA1990.DAT.*
12: /SID52/gcgdata/geneseq/geneseq/NA1991.DAT.*
13: /SID52/gcgdata/geneseq/geneseq/NA1992.DAT.*
14: /SID52/gcgdata/geneseq/geneseq/NA1993.DAT.*
15: /SID52/gcgdata/geneseq/geneseq/NA1994.DAT.*
16: /SID52/gcgdata/geneseq/geneseq/NA1995.DAT.*
17: /SID52/gcgdata/geneseq/geneseq/NA1996.DAT.*
18: /SID52/gcgdata/geneseq/geneseq/NA1997.DAT.*
19: /SID52/gcgdata/geneseq/geneseq/NA1998.DAT.*
20: /SID52/gcgdata/geneseq/geneseq/NA1999.DAT.*
21: /SID52/gcgdata/geneseq/geneseq/NA2000.DAT.*
22: /SID52/gcgdata/geneseq/geneseq/NA2001.DAT.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2125	100.0	2125	19	Humanised vector p
2	1203	56.6	1425	19	Humanised vector p
3	1194.4	56.2	1911	19	Humanised vector p
4	671.2	31.6	4299	14	Sequence encoding
5	671.2	31.6	4473	20	HER-2 nucleic acid
6	670.4	31.5	9274	22	HER2 transgene pla
7	669.6	31.5	4530	16	Her-2/neu (ERBB2)/c
8	669.6	31.5	4530	18	Human HER2 gene.
9	669.6	31.5	4530	21	Nucleotide sequenc
10	669	31.5	4472	21	cDNA encoding the
11	667.6	31.4	3768	17	HER-2/neu oncogene

12	667.6	31.4	3768	20	AAO1912	Human HER-2/neu on
13	667.6	31.4	3768	22	AAH23392	Human HER-2/neu pr
14	666	31.3	3768	21	AAA09455	Human heregulin 2
15	508.8	23.9	3771	21	AAH9737	Mouse Her-2/neu CD
16	508.8	23.9	3771	22	AAH42210	Nucleotide sequenc
17	499.6	23.5	3600	21	AAH89736	Human HER-2/neu co
18	495.8	23.3	3955	16	AAO1590	Rat neu promoter.
19	495.8	23.3	3955	21	AAH89753	Rat HER-2/neu prot
20	474.8	22.3	1115	21	AAF21778	Human breast and o
21	459.8	21.6	3553	15	AAQ74208	Expression vector
22	454.8	21.4	7073	18	AAH50962	TF8-SG9 CDR-grafte
23	454.8	21.4	7864	18	AAH50963	TF8-SG9 CDR-grafte
24	454.8	21.4	11795	21	AAA59345	Nucleotide sequenc
25	454.8	21.4	13254	17	AAH40915	Nucleotide sequenc
26	454.8	21.4	13254	21	AAA32151	p12 Combo BM 12
27	454.8	21.4	13254	21	AAA32165	Complete nucleotid
28	453.8	21.4	608	19	AAV21731	ColEI origin of re
29	453.8	21.4	5241	17	AAV217307	pHL104 containing
30	453.8	21.4	5365	10	AAH90646	Nucleotide sequenc
31	453.8	21.4	5413	10	AAH90649	Nucleotide sequenc
32	453.8	21.4	5518	10	AAH90647	Nucleotide sequenc
33	453.8	21.4	5566	10	AAH90648	Nucleotide sequenc
34	453.8	21.4	5579	11	AAQ02032	Plasmid pHS 164.
35	453.8	21.4	6051	12	AAQ14934	pBG391 sequence.
36	453.8	21.4	6149	10	AAH90645	Sequence of T4 lym
37	453.8	21.4	6151	10	AAH90644	Sequence of T4 lym
38	453.8	21.4	6151	11	AAQ03006	Full length T4 CDN
39	453.8	21.4	6151	11	AAQ05608	Plasmid pBG381 inc
40	453.8	21.4	6165	13	AAQ20324	Sol.rhesus CD4 seq
41	453.8	21.4	6477	11	AAQ20330	Plasmid pinf 4-49.
42	453.8	21.4	7252	20	AAH86928	Complete sequence
43	453.8	21.4	7377	11	AAQ03005	Full length T4 CDN
44	453.8	21.4	7377	11	AAQ05607	Plasmid p170-2 inc
45	453.8	21.4	8119	17	AAH33477	Subtilisin N62D/G1

ALIGNMENTS

RESULT	1
AAV21727	AAV21727 standard; cDNA; 2125 BP.
ID	AAV21727 standard; cDNA; 2125 BP.
XX	AC AAV21727;
XX	DT 17-AUG-1998 (first entry)
XX	Humanised vector pITL-hHER/neu.
DE	Vector; vaccine; tumour; antigen; plasmid pITL-hHER/neu;
XX	human; HER-2/neu; C-erbB-2; breast cancer; ds.
KW	Chimeric - Homo sapiens.
XX	Chimeric - Escherichia coli.
OS	Synthetic.
XX	Key
XX	CDS
XX	Location/Qualifiers
FT	13..921
FT	/tag= a
FT	/product= human HER-2/neu
FT	922..1181
FT	/tag= b
FT	/note= "combined splice and polyA sequences"
FT	1195..1401
FT	/tag= c
FT	/note= "SupF gene"
FT	1412..1864
FT	/tag= d
FT	/note= "ColEI origin of replication"
FT	1877..2125
FT	/tag= e
FT	/note= "RANTES promoter"
XX	

```
PN W09806863-A1.
XX
PD 19-FEB-1998.
XX
PF 14-AUG-1997; 97WO-US14306.
XX
PR 14-AUG-1996; 96US-0023931.
XX
PA (USSH ) US DEPT HEALTH & HUMAN SERVICES.
XX
PI Nelson EL, Nelson PJ;
XX
XX WPI; 1998-159552/14.
XX
DR Humanised polynucleotide vectors - comprising human derived promoter
XX PT and sequence acceptance site, used for the production of vaccines
XX
XX Example 7; Page 41-42; 125pp; English.
XX
CC Plasmid pITL-hHER2/neu comprises base vector pITL (see AAV21724) and
CC a human HER-2/neu nucleic sequence. pITL-hHER2/neu was used to
CC evaluate the toxicity of anti-tumour vaccination in rats, and in
CC phase I and phase II trials to evaluate polynucleotide vaccination
CC in advanced breast cancer. Novel humanised vectors, which can be
CC based on pITL, comprise a human-derived promoter or mammalian
CC homologue which is functional in mammalian target tissue and cells
CC and a sequence acceptance site which accepts cDNA products from
CC RT-PCR cloning. The vectors are non-replicating in mammalian cells
CC but are capable of extended stable expression of the target
CC sequence, generating an immune response in immunised individuals.
CC The vectors selectively elicit immune responses to the target
CC sequences with little or no immune response to the other components
CC of the vectors.
XX
SQ Sequence 2125 BP; 449 A; 550 C; 598 G; 428 T; 0 other;

Query Match 100.0%; Score 2125; DB 19; Length 2125;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 2125; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GCCACATGGCCCTGACCTCTCTACATGCCATCTGGAATTTCCAGATGAGAGGGC 60
Db 1 gccaccatggccctgaacctctctacatgccatctggaatttccagatgagaggc 60
Qy 61 GCATGCCAGCTTGCCCATCACTGACCCACCTCTGTGGACCTGGATGACAAGGC 120
Db 61 gcatgccagcttgcccatcaactgacccactctctgtgtggacctggaagaggc 120
Qy 121 TGCCCGCGAGCAGAGCCAGCCCTCTGACGTCCATCATCTCTGCGGTGGTGGCAAT 180
Db 121 tgcccgcgagcagagaccagccctctgacgtccatcatctctgsggtggcaat 180
Qy 181 CTGCTGGTGTGCTTGGGGGTGGTCTTTTGGGATCTTCATCAAGCGACGCGAGCAAG 240
Db 181 ctgctgggtgtgtgtgggtgtgttttggatccctcatcaagcgacgagcagaag 240
Qy 241 ATCATGTCCAGACCTGCCCCGGCGCTGGGGCATGGTCCACACAGCCAGCCGACG 300
Db 241 atcatgtccagacctgccccggcgctggggcatgggtccaccacagcaccgcagc 300
Qy 301 TCATCTACAGAGTGGCGGTGGGACCTGACACTAGGCGTGGAGCCCTCTGAAGAGGAG 360
Db 301 tcactaccagagtggcggtgggacctgacctgacctgacctgacctgacctgacct 360
Qy 361 GCGCCAGGTCTCCATGCGACACCTCCGAAGGGGTGCGTCCGATGATTTGATGGTGAC 420
Db 361 gccccaggtctccactggcaccctccgaagggtggtgctccgatgtatttgggtgac 420
Qy 421 CTGGGATGGGGGAGCCAGGGCTGCAAGAGCTCCCCACACATGACCCGAGCCCTCTA 480
Db 421 ctgggaatgggggagcccaagggtgctgcaaggctccccacacatgacccccagccctcta 480

481 CAGCGGTACAGTGAAGACCCACAGTACCCCTGCGCTGTGAGCTGATGGCTACGTTGCC 540
481 cagcggtacagtgaagacccacacagtagtaccctgacctgagactgagtgactgtgcc 540
541 CCCGTGACCTGCAGCCCCCAGCCTGAATATGTGAACAGCCAGATGTTCCGCCAGCC 600
541 cccgtgacctgcagccccagcctgaatgtgaaccagccagatgttccggccccagccc 600
601 CTTTCGCCCGAGAGGGCCCTCTGCTGCTGCCACCTGCTGTGTCGCACCTCTGAAAAG 660
601 ctttcgcccgagagggccctctgctgctgcgcagacctgctggtgcaactctgaaaag 660
661 CCCAAGACTCTCTCCCGAGGAAGAAATGGGTGCTGCTAAAGACGTTTTTGGCTTGGGGGT 720
661 cccaagactctctcccaggaagaatgggtcgctcaagacgtttttgctttgggggt 720
721 GCGGTGAGAACCCGAGACTTGACACCCAGGAGGAGCTGCCCTCAGCCGCCACCCCTC 780
721 gccgtgagaaccccgagacttgacaccccgaggaggtgagctgccccctcagccccccctc 780
781 CTCCTGCTTCAGCCCCAGCCTTCGACACCTCTATTACTGGACCCAGACCCACACAGAGC 840
781 ctctgcttcagccccagccttcgacaacctctattactggaccagaccaccagagc 840
841 GGGGGGTCTCCACCGACACTTCAAGGGACACCTACCGGACAGAACCCAGTACCTGG 900
841 ggggggtctccaccagaccttcaaggagcacctacgagcagagaaacccagagctctgg 900
901 GTCTGGAGTCCCGAGCTGAAGCCTTAAGGCCATATGCTGAGTGGATGCTTGACCCCA 960
901 gtctggagtcctcgagctggaagccttaaggcccatatggtgagtgatgaccttgacccca 960
961 GCGGGGATGGGGAGACCTGTAGTCAGAGCCCGCGGACGACACAGGCCAATGCCCGTCC 1020
961 ggcgggatggggagacctgtagtcagagcccgcgagcagcccgagcagacagccaatgccc 1020
1021 TTCCCTGCACTGAGTAGTACTGCCCCGGGTGGATCCTGTGACCCCTCCCGAGTGCT 1080
1021 ttccctgcagtgagttagtgcctgggtgggtacctctgtgacccctcccgagtgct 1080
1081 CTCCTGGCCCTGGAAGTTGCCACTCCAGTGCCACACGCTTGTCTAATAAATAAGT 1140
1081 ctctggccctgggaagtggccactccagtgccaccagcctgtctctataataaataagt 1140
1141 TGCATCATTTTGTCTGACTAGTGTCTCTATAATATTATTAAGTTGATATCGAATTTT 1200
1141 tgcatactatttgtctgactagtgctctataattattataagcttgatgatactgaattct 1200
1201 TCTCAACGTAACTTTACAGCGGCGCTCATTTGATGATGATGCGCCCGCTTCCCGATA 1260
1201 tctcaacgtaacctttacagcgcgctcatttgatgatgagcgcccgcttcccgata 1260
1261 AGGAGCAGGCGCAGTAAAGCATTAACCGTGGTGGGTTCGCGAGCGCCAAAGGAGCA 1320
1261 aggagcagcgccagtataaagcattaccctggtgggttccccgagcgcccaaaaggagca 1320
1321 GACTCTAATCTGCGCTCATCGACTTCGAAGGTTTCCGAATCTTCCCGCCACCACTCACT 1380
1321 gactctaatctgctgcgtcatcgacttcgaagtttccaatcttccccaccacatcaact 1380
1381 TTCAAAAGTCGAAGAATTCCTGAGCCCGGTGATGCGGTAGCCGTAGTAGGCCACCACTCAAG 1440
1381 ttcaaaagtcgaaagaattctcgagcccgtagcgctgtagcgctgtaggcacacactcaag 1440
1441 AACCTGTAGCACCGCTACATACCTGCTGCTTAATCTGTTACCACTGGCTGCTGCC 1500
1441 aacctgtagcacccgctacataacctgctgctgctaaacctgtaccagtggtgctgcc 1500
1501 AGTGCGCATAGTGTGTCTTACCGGGTTGACTCAAGACGATAGTTACCGGATAAGGCG 1560
1501 agtgcgataagtgctgtcttaccgggttgactcaagacgatatgtaccggataaggcg 1560
1561 CAGCGTGGCGCTGAACGGGGGTTCTGTGCACACAGCCCGCTTGGAGGACGACCTAC 1620
```

```
Db 1561 cagcggctcgggctgaacggggggttcgtagacacagccagcttgagcgaacgacctac 1620
Qy 1621 ACCGACTGAGTACTACTACAGCGTGAGCATTTGAGAAAGCGCCACCGCTTCCCGAAGGGAGA 1680
Db 1621 accgaactgagatacctacagcgtgagcattgagaaagcgcacgcttcccgaaggagga 1680
Qy 1681 AAGCGGACAGGTATCCGTAAGCGGCGAGGTCGCAACAGGAGAGCGCACGAGGAGGCTT 1740
Db 1681 aagcggaacaggtaaccggtgaagcgaggggtcggaacaggagagcgcaagggagcctt 1740
Qy 1741 CCAGGGGAAACCGCTGGTATCTTTATAGTCTGTCGGGTTTTCGCCACCTCTGACTTGAG 1800
Db 1741 ccaggggaaacgcgctgtatctttatagtcctgctcgggtttcgccacctctgacttgag 1800
Qy 1801 CGTCGATTTTGTGATGCTGTCGAGGGGGGCGAGCCTATGGAATAACGCCAGCAACGGG 1860
Db 1801 cgtcgattttgtgatgctcgtcagggggcgagcctatggaataacgccagcaacgcg 1860
Qy 1861 GCCGGGATCCGGAGAGTCACTCTAGATGAGAGAGCAGTGAGGAGAGACAGAGACTC 1920
Db 1861 gccggggatccggagagcctcactctagatgagagcagtgagggagagagagacctc 1920
Qy 1921 GAATTTCCGGAGCTATTTCACTTTTCTTTCCGTTTGTGCAATTTCACTTATGATACCG 1980
Db 1921 gaatttcggagcatttcagtttcttcttcctgtttgtgcaatttcacttatgataccg 1980
Qy 1981 GCCATGCTGGTGGCTATTTTGGAAACTCCCTTATAGGGATGCCCCCTCACTTGGCCCTA 2040
Db 1981 gccaatgctgtgtgctattttggaaactcccccttaggggatgcccccaactgagcccta 2040
Qy 2041 TAAAGGGCCAGCTCAGCTGCGAGAGGATTCCTGCGAGAGATCAAGACAGCAGCTGACCT 2100
Db 2041 taaaggccagcctgagctgcagagagattccctgcagagatcaagacagcagctggacct 2100
Qy 2101 CGCACAGCTCTCCACAGGTACCT 2125
Db 2101 cgcacagcctctccacaggtacct 2125

RESULT 2
AAV21724
ID AAV21724 standard; cDNA; 1425 BP.
AC AAV21724;
XX
XX 17-AUG-1998 (first entry)
DE Humanised vector pITL.
XX
XX Vector; vaccine; tumour; antigen; plasmid pITL; ds.
XX
OS Chimeric - Homo sapiens.
OS Chimeric - Escherichia coli.
OS Synthetic.
FH Key Location/Qualifiers
FT misc_feature 1..221 /*tag= a
FT /*note= "stuffer sequence"
FT polyA_site 22..481 /*tag= b
FT /*note= "combined splice and polyA sequences"
FT CDS 495..701 /*tag= c
FT /*note= "SupF gene"
FT misc_feature 712..1164 /*tag= d
FT /*note= "ColE1 origin of replication"
FT promoter 1177..1425 /*tag= e
FT /*note= "RANTES promoter"
```

```
PN WO9806863-A1.
XX
XX 19-FEB-1998.
XX
XX 14-AUG-1997; 97WO-US14306.
XX
XX 14-AUG-1996; 96US-0023931.
XX
XX (USSH ) US DEPT HEALTH & HUMAN SERVICES.
XX
XX Nelson EL, Nelson PJ;
XX
XX WPI; 1998-159552/14.
XX
XX Humanised polynucleotide vectors - comprising human derived promoter
XX and sequence acceptance site, used for the production of vaccines
XX
XX Claim 14; Page 29-30; 125pp; English.
XX
XX Plasmid pITL comprises a base vector for novel humanised
XX polynucleotide vectors. Such vectors comprise a human-derived
XX promoter or mammalian homologue which is functional in mammalian
XX target tissue and cells and a sequence acceptance site (see
XX AAV21735-36), which accepts cDNA products from RT-PCR cloning. They
XX also contain minimal non-human components, such as a replication
XX origin (see AAV21715) and selectable marker gene (see AAV21717-18) that
XX are necessary for production of the vector, as well as human-derived
XX splice and polyA sequences (see AAV21723). The novel vectors are
XX used to express target antigens, especially tumour antigens. They
XX are non-replicating in mammalian cells but are capable of extended
XX stable expression of target sequences generating an immune response
XX in immunised individuals. The vectors selectively elicit immune
XX responses to the target sequences with little or no immune response
XX to the other components of the vectors. The target antigens are
XX expressed as intracellular polypeptides or peptides and, as such,
XX are processed as self polypeptides or peptides and appropriately
XX presented on antigen presenting cells.
XX
XX Sequence 1425 BP; 325 A; 381 C; 404 G; 315 T; 0 other;
```

```
Query Match 56.6%; Score 1203; DB 19; Length 1425;
Best Local Similarity 100.0%; Pred. No. 2.9e-299;
Matches 1203; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 922 GCCTTAAGGGCCATATGCTGAGTGATGCTTGACCCCGAGGGGATGGGGAGACCTG 981
Db 223 gccttaaggggccatatggtgagtgagcttgaccccgaggggatggggagacctg 282
Qy 982 TAGTCAGAGCCCGGGCAGCAGCAGCCCAATGCCCTTCCCTGCAGTGAGTAGTGA 1041
Db 283 tagtcagagcccccggggcagcagagggccaatgccctccctcgcagtgagtagtga 342
Qy 1042 CTGCCCGGGTGGGATCCCTGTGACCCCTCCCGAGTGCCTCTCTCGCCCTGGGAAGTTGCC 1101
Db 343 ctgcccgggtgggatccctgtgacccctcccgagtcctcctcctcctggcgaagtgc 402
Qy 1102 ACTCCAGTCCCGCAGCCCTTGTCTTAATAAATTAAAGTTGCATCATTTTGTCTGACTAG 1161
Db 403 actccagtgcccaccagcctgttcctaaataaataaagttgcatcattttgtctgactag 462
Qy 1162 GTGTCCTCTATATATTATAAGCTTGTATCGAATTTCTTCAAGTACACTTTACAG 1221
Db 463 gtgtcctctataataataagccttgatcgaaattcttctcaacgtaacactttacag 522
Qy 1222 CGGCGCGTCATTTGATATGATGCGCCCGCTTCCCGATAAGGGAGCAGGCCAGTAAAGC 1281
Db 523 cggcgctcatttgatgatgatgcgcccgcttcccataagggagcagggccagtaaaagc 582
Qy 1282 ATTACCCGTTGGTGGGTTCCCGAGCGGCCAAAGGGAGCAGACTCTTAATGTCCGCTCATC 1341
Db 583 attacccgtggtggggttcccagagcgccaaaggagcagactcctaaattctgcgtcatc 642
```


||||| 836 ggtggatccctgtgaccctccctcctcctcctgacctggaagtgcacctccag 895
QY 1109 TGCCACCACCGCTTCCCTAAATAAATAAGTTGTCATCATTTTGTCTGACTAGGTGTCCT 1168
Db 896 tgcaccacagctgttcccaataaataaagttgctcatcttttctgactaggttccct 955
QY 1169 CTATAATATATAAGCTTGATATCGAATCTTCTCAACGTAACTTTACAGCGCGCG 1228
Db 956 ctataattataaagcttgatcgaattcttctcaacgttaacactttacagcgcg 1015
QY 1229 TCATTTGATATGATGCGCCCGCTTCCGATAAAGGAGCAGGCCAGTAAAGCATTTACCC 1288
Db 1016 tcatttgatgatggtgcccccttccgataaaggagcagcgccagtaaaagcattacc 1075
QY 1289 GTGGTGGGTTCGCGAGCGGCAAGGAGCAGACTCTTAATCTGCGTCATCGACTTCG 1348
Db 1076 gtggtggtgtcccgagcgccaaaggagcagactctaaatctgcgtcatcgacttccg 1135
QY 1349 AAGTTTCGAATCTTCCCCACCACCATCTTCAAAAGTCCGAAAGAAATTCCTGCAGC 1408
Db 1136 aaggttcgaatcctccccaccaccatcaacttcaaaagtcgaaagaattcctgcagc 1195
QY 1409 CCGTGTAGCGGTAGTTAGGCCACCATCTCAAGAACTCTGTAGCACCGCCTTACATACCTCG 1468
Db 1196 ccgtgtagccgttagtgccaccacttcaagaactctgtagcaccgctacatacctcg 1255
QY 1469 CTCGTCTAATCTTGTACAGTGGCTGCTGCCAGTGGCGATAAGTGTCTTTACCGGGT 1528
Db 1256 ctctgtaactctgttaccagtggtcgtcgcaagtggcgataaagtgcttaccgggt 1315
QY 1529 TGGACTCAAGACGATAGTTACCGGATAAAGCGCAGCGGTGCGGCTGAACGGGGGTTCGT 1588
Db 1316 tgaactcaagacatagttaccgataaaggcagcggtcggtgtaaaagggttccgt 1375
QY 1589 GCACACGCCACTTGGAGCGAAGCACTTACACCGAACTGTAGATACCTTACACGCTGAGC 1648
Db 1376 gcacacagccagcttggagcgagcactacaccgaactgagatacctacagcgtgagc 1435
QY 1649 ATTGAGAAAGCGCCAGCTTCCGAGAGGAGAAAGCGGACAGGTATCCGGTAAGCGGCA 1708
Db 1436 attgagaaagcgccacgcttcccgagggagaaagcgagaggtatccggtaagcgca 1495
QY 1709 GGCTCGGAACAGAGAGCGCACGAGGAGGCTTCCAGGGGAAACGGCTGTATCTTTATA 1768
Db 1496 ggtcggaacagagagcagcagagagagcttccagggggaaagcgtggtatcttata 1555
QY 1769 GTCCGTGCGGTTTCCGACCTCTGACTTGAGCGTGCATTTTGTGATGCTCTCAGGGG 1828
Db 1556 gtccgtgcgggttccgacctgtgactgagcgtcgatatttctgctgctcagggg 1615
QY 1829 GCGGAGCCTATGGAAGAGCCAGCAACGCGCGCGGGGATCCGGAGAGCTCAGTCTAG 1888
Db 1616 ggcgagcctatggaaaaacgcagcaacgcgagcgccggggatccggagagctcactag 1675
QY 1889 ATGAGAGAGCAGTGGAGGAGAGACAGAGACTCGAAATTTCCGGAGCTATTTCAGTTTCTT 1948
Db 1676 atgagagacagtggagggagagacagagactgaatttccggagctatttcaagtttctt 1735
QY 1949 TTCGGTTTGTGCAATTTCACTTATGATACCGGCCAATGCTTGGTGTCTATTTTGAAAC 2008
Db 1736 ttccgttttgtgcaatttcaacttatgataccggccaaatgctgtgttatttttggaaac 1795
QY 2009 TCCCTTAGGGGATGCCCTCACTGCTTAAAGGGCCAGCGTGTAGTGAGAGGAT 2068
Db 1796 tcccttaggggagtccttcaactgagcctataaagggcccagcctgagctgagaggtat 1855
QY 2069 TCTGTGAGAGGATCAAGACAGCAGTGGACCTCGCACAGGCTCTCCACAGGTACC 2124
Db 1856 tctgtcagagga tcaagacagacagctggacctcgcagcagcctctccacaggtacc 1911

RESULT 4

AAQ46083
ID AAQ46083 standard; cDNA; 4299 BP.
XX
-AC AAQ46083;
XX
DT 07-FEB-1994 (first entry)
XX
DE Sequence encoding a c-erbB-2 tumour antigen.
XX
KW Tumour antigen; c-erbB-2; glycoprotein; ss.
XX
OS Homo sapiens.
XX
FH Key Location/Qualifiers
FT CDS 1..4299
FT- /*tag= a
XX
PN WO9316185-A.
XX
PD 19-AUG-1993.
XX
PF 05-FEB-1993; 93WO-US01055.
XX
PR 06-FEB-1992; 92US-0831967.
XX
PA (CETU) CETUS ONCOLOGY CORP.
PA (CREA-) CREATIVE BIOMOLECULES INC.
XX
PI Houston LL, Huston JS, Oppermann H, Ring DB;
DR WPI; 1993-272889/34.
DR P-PSDB; AAR39568.
XX
PT New single chain Fv polypeptide binding to C-erbB-2 tumour
antigen - for imaging or treating breast or ovarian cancer etc.
XX
PS Disclosure; pages 48-54; 87pp; English.
XX
CC c-erbB-2 refers to a protein antigen expressed on the surface of
CC tumour cells, such as breast and ovarian tumour cells, which is an
CC approx. 200,000 mol. wt. acidic glycoprotein having an isoelectric
CC pt. of about 5.3 (see AAQ46083, AAR39568).
XX
SQ Sequence 4299 BP; 882 A; 1307 C; 1266 G; 844 T; 0 other;

Query Match 31.6%; Score 671.2; DB 14; Length 4299;
Best Local Similarity 95.4%; Pred. No. 1.7e-162;
Matches 702; Conservative 0; Mismatches 33; Indels 1; Gaps 1;

QY 228 ACGGACGACGAAGATCACATGTCCAGACCTGCCGCGCGCTGGGGCATGTGCCACCA 287
Db 3075 accccagcagggcttctctgtccagacctgcccgggcgtggtgggcatggtccacca 3134
QY 288 CAGGCACCGCAGCTCATCTACAGGAGTGCGGTGGGGACCTGACACTAGGCTGGAGCC 347
Db 3135 caggcaccgagctcatctaccagagtgcggtgggacctgacactgagcgtgagcc 3194
QY 348 CTCTGAAGAGGAGGCCCCCAGGTCTCCACTGGCACCTCCGAAGGGGCTGGCTCCGATGT 407
Db 3195 ctctgaagagagggccccaggtctccactggcacctccgaagggtggtccgatgt 3254
QY 408 ATTTGATGTTGACTGGGATGGGGCAGCAAGGGGCTCAAGGCTCCCCACACATGA 467
Db 3255 atttgatggtgacctgggaatgggagccaaagggtctcaaaagcctcccacacatga 3314
QY 468 CCCAGCCCTCTACAGCGGTACAGTACAGGACCCACAGTACCCCTGCCCCTCTGAGACTGA 527
Db 3315 cccagccctctacagcgttacagtgaggaccaccaagtagtaccctctctgagactga 3374
QY 528 TGCTACGTTGCCCTGACCTGCAGCCCCCAGCCCTGGAATATGTGAACAGCCAGATGT 587
Db 3375 tggctacgttggccccctgacctgcagccccccagcctgaatatgtgaaccagcagatgt 3434

QY 588 TCGGCCCCAGCCCTTGGCCCGAGAGGCGCCTTGCCTGCTGCTCCGACCTGCTGTCG 647
 Db 3435 tegggccagccccccttcgcccagagagggccctctgctgctgcccagcctgctgctc 3494
 QY 648 CACTCTGGAAGAGCCCAAGACTCTCTCCCGAGGAGAGATGGGTCGTCGAAGACGTTT 707
 Db 3495 cactctggaaagcccaadactctctcccagggaaagatgggtcgtcaagacgctttt 3554
 QY 708 TGCTTTGGGGTGGCTGGAGAACCCGAG-ACTTGACACCCCGAGGAGGAGCTGCCCC 766
 Db 3555 tgcctttgggggtgccttgagaaaccccgagtagcttgacacccagggagggctgcccc 3614
 QY 767 TCAGCCCCACCTCTCTGCTGCTTACAGCCAGCCTTCGACAACTCTATTACTGGGACCA 826
 Db 3615 tcagccccacctctctctgcttgaagccagccttcgacaaacctattactggacca 3674
 QY 827 GGACCCACAGAGCGGGGGCTCCACGACGACCTTCAAGAGGACACCTACGCGAGAA 886
 Db 3675 ggaacccagagagcggggggtccacccagacaccttcaaaaggagacacctacggcagaa 3734
 QY 887 CCAGAGTACCTGGGTCTGAGCTGCGAGTGTGAAGCCCTTAAGGGCCATATGCTGAGTGG 946
 Db 3735 ccagagtagctgggtctgagctgccaagtgaaccagaaagccaagtcgcagaagcc 3794
 QY 947 ATGCTTGACCCGAGG 962
 Db 3795 ctgatgtctctcagg 3810

RESULT 5
 AAZ31071
 ID AAZ31071 standard; DNA; 4473 BP.
 XX AAZ31071;
 XX 17-JAN-2000 (first entry)
 DE HER-2 nucleic acid sequence.
 XX HER-2; c-neu; ErbB2; transmembrane receptor; tyrosine kinase activity;
 KW epidermal growth factor receptor; EGFR; HER-1; cancer; breast cancer;
 KW ovarian cancer; gastric cancer; antisense oligonucleotide; expression;
 KW hyperproliferative disease; ss.
 XX Homo sapiens.
 OS
 XX W09948906-A1.
 PN
 XX 30-SEP-1999.
 PD
 XX 25-MAR-1999; 99WO-US06492.
 PF
 XX 26-MAR-1998; 98US-0048804.
 PR
 XX (ISIS-) ISIS PHARM INC.
 PA (PENN-) PENN STATE RES FOUND.
 XX
 XX Bennett CF, Lipton A, Witters LM;
 PI WPI: 1999-610749/52.
 XX
 XX New antisense sequences used to treat hyperproliferative conditions,
 PT especially cancer -
 XX
 XX Examples: Page 38-39; 44pp; English.
 PS
 XX This is the human HER-2 polynucleotide sequence. The HER-2 gene also
 CC called c-neu and ErbB2, encodes a transmembrane receptor, with tyrosine
 CC kinase activity. HER-2 is related to the epidermal growth factor
 CC receptor (EGFR or HER-1). Aberrant HER-2 expression is present in a wide
 CC number of cancers, especially breast, ovarian and gastric cancers. This
 CC sequence is used in the invention to design 12-25 nucleotide

CC oligonucleotides that decrease the expression of human HER-2. The
 CC oligonucleotides of the invention (AAZ31067-Z31070) can also be used for
 CC modulating the expression of human epidermal growth factor receptor. The
 CC oligonucleotides are used to treat diseases or conditions associated with
 CC HER-2, particularly hyperproliferative diseases such as cancer.
 XX
 SQ Sequence 4473 BP; 902 A; 1383 C; 1329 G; 859 T; 0 other;
 Query Match 31.6%; Score 671.2; DB 20; Length 4473;
 Best Local Similarity 95.4%; Pred. No. 1.8e-162;
 Matches 702; Conservative 0; Mismatches 33; Indels 1; Gaps 1;
 QY 228 ACCGCGACAGAGATGACATGTCACAGACCTGCCCGGCGCTGGGGCATGCTGCACCA 287
 Db 3249 acccagcagggctctctctgtccagacccctgccccggcgctgggcatggtccacca 3308
 QY 288 CAGGACCGCAGCTCATCTACAGAGTGGGGTGGGACCTGACACTAGGGCTGGAGCC 347
 Db 3309 caggcaccgcagctcatctaccagagtggggtgggacctgacactagggtgagcc 3368
 QY 348 CTCTGAAGAGAGGCCCGCCAGGTCTCCACTGCGACCCCTCCGAAGGGGCTGCTCCGATGT 407
 Db 3369 cctctgaagagagggccccaggtctccactggcacctccgaagggtggtcgcagatgt 3428
 QY 408 ATTTGATGTTGACCTGGGAATGGGGGCGAGCCAAAGGGGCTGCAAAAGCCCTCCACACATGA 467
 Db 3429 atttgatgtgacctgggaatggggcagccaaagggtgcaaaagctccccacacatga 3488
 QY 468 CCCAGAGCCCTCTACAGCGGTACAGTGGAGAGCCGCCACAGTACCCCTGCGCTGTGAGACTGA 527
 Db 3489 cccagagccctctacagcggtacagtggagagccacacagtacccctgcccctctgagactga 3548
 QY 528 TGGCTACGTTGCCCGCTGACCTGACGCGCCCGAGGCGCCTTGCCTGCTGCGCCGACCTGCTGTCG 647
 Db 3549 tggctacgttgcctccctgacctgagccccagcctggaatgtgacccagccagatgt 3608
 QY 588 TCGGCCCCAGCCCTTGCCTGCGAGAGGCGCCTTGCCTGCTGCGCCGACCTGCTGTCGTCG 647
 Db 3609 tcggccccagcccccttcgccccgagagggccctctgctgctgcccagcctgctggtgc 3668
 QY 648 CACTCTGAAGAGCCCGACACTCTCTCCCGAGGAGAGATGGGTCGTCGAAGAGCTTTT 707
 Db 3669 cactctggaagggcccaagactctctcccagggaaagatgggtcgtcaagacgctttt 3728
 QY 708 TGCCTTTGGGGTGGCTGGAGAACCCGAG-ACTTGACACCCCGAGGAGGAGCTGCCCC 766
 Db 3729 tgcctttgggggtgccttgagaaaccccgagtagcttgacacccagggagaggtgcccc 3788
 QY 767 TCAGCCCCACCTCTCTGCTGCTTACAGCCAGCCTTGCACAACTCTATTACTGGGACCA 826
 Db 3789 tcagccccacctctctctgcttgaagccagccttcgacaaacctctattactggacca 3848
 QY 827 GGACCCACAGAGCGGGGCTCCACGACGACCTTCAAGAGGACACCTACGCGGAGAA 886
 Db 3849 ggaacccagagcggggggtccacccagacaccttcaaaaggagacacctacggcagaa 3908
 QY 887 CCCAGAGTACCTGGGTCTGAGCTGCGAGTGTGAAGCCCTTAAGGGCCATATGCTGAGTGG 946
 Db 3909 ccagagtagctgggtctgagctgccaagtgaaccagaaagccaagtcgcagaagcc 3968
 QY 947 ATGCTTGACCCGAGG 962
 Db 3969 ctgatgtgctctcagg 3984

RESULT 6
 AAF24297
 ID AAF24297 standard; DNA; 9274 BP.
 XX
 AC AAF24297;
 XX
 DT 03-APR-2001 (first entry)

Db 3525 tggctacgttgcgcctgaactgagcccccagcgtgaatatgtgaaccagccagatgt 3584
Qy 588 TCGGCCCCAGCCCTTGCCTCCGAGAGGGCCCTCTGCTGCTGCCGACCTGCTGGTGC 647
Db 3585 tcggcccccagcccttcgcccagagaggccctctgctgctgcccagcctgtggtgc 3644
Qy 648 CACTCTGAAAGGCCCAAGACTCTCTCCCGAGGGAAGTGGGTGCTCAAGAGCTTTT 707
Db 3645 cactctgaaaggccaaagactctctcccagggaagaatgggtcgcaaaagcgtttt 3704
Qy 708 TGCCTTTGGGGTGCCTGGAGAACCCCGAG-ACCTGACACCCCGAGGAGAGTGCCTCC 766
Db 3705 tgcctttgggggtgcgtggagaaaccccgagtacttgacaccccgaggagcgtgcctc 3764
Qy 767 TCAGCCCCACCTCTCTGCTTCAGCCCGAGCTTCGACACCTTACTTACTGGACCA 826
Db 3765 tcagcccccacccctctctgcttcagcccgagccttcgacacccctctattactggacca 3824
Qy 827 GGACCCACAGAGCGGGGGTCCACCCAGCACCTTCAAAAGGACACCTAGGGCAGAAA 886
Db 3825 ggaccaccagagcgggggtccaccagacacttcaaaaggacacactagcgagagaa 3884
Qy 887 CCCAGAGTACCTGGGTCTGGAGCTGCCAGTGTGAAGCCCTTAAGGGCCATATAGTGG 946
Db 3885 cccagagtacctgggtctggagctgcccagtgtaaccagaaagcccaagtcgcgagaagcc 3944
Qy 947 ATGCTTGACCCACAGG 962
Db 3945 ctgatgtgtcctcagg 3960

RESULT 9

AAZ60815
ID AAZ60815 standard; DNA; 4530 BP.

AC AAZ60815;

XX 16-MAY-2000 (first entry)

XX Nucleotide sequence of a cognate transgene of c-neu.

XX Cognate transgene; CTG; tumourigenic; cellular immunogen; immunisation;
KW proto-oncogene; malignancy; allogenic cell; vaccine; cancer; ss.

XX Homo sapiens.

XX WO200004927-A1.

XX 03-FEB-2000.

XX 08-JUL-1999; 99WO-US15594.

XX 24-JUL-1998; 98US-0093965.

XX (UYAL-) UNIV ALLEGHNEY HEALTH SCI.

XX (HALP/) HALPERN M S.

XX (ENGL/) ENGLAND J M.

XX Halpern MS, England JM;

XX WPI; 2000-182543/16.

XX Cellular immunogens comprising allogenic donor cells transfected with a
PT construct comprising a proto-oncogene cognate, useful as cancer
PT vaccines

XX Disclosure; Page 66-68; 77pp; English.

XX The present sequence represents a cognate transgene (CTG) which is
CC rendered non-tumourigenic by deletion of amino acids.1-731. The CTG
CC is used in the course of the invention. The specification describes a
CC cellular immunogen for immunizing a host against the effects of the

CC product of a target proto-oncogene which is associated with a
CC malignancy. The cellular immunogen comprises allogenic cells
CC transfected with transgene construct comprising a transgene cognate
CC to target proto-oncogene and a strong promoter. The cellular immunogen
CC is useful for vaccinating a host against cancer by inserting the
CC transgene construct into the body of the host for the expression of
CC the transgene. The method of the invention is designed to target
CC mutation-driven non-self determinants. The cellular immunogens induce
CC reactivity for self-determinants in the over expressed product of tumour
CC associated and over expressed proto-oncogenes.
XX
SQ Sequence 4530 BP; 922 A; 1382 C; 1346 G; 880 T; 0 other;

Query Match 31.5%; Score 669.6; DB 21; Length 4530;

Best Local Similarity 95.2%; Pred. No: 4.5e-162;

Matches 701; Conservative 0; Mismatches 34; Indels 1; Gaps 1;

Qy 228 AGCGCAGCAGAAATCACATGTCCAGACCTCGCCCGGCGCTGGGGCGATGGTCCACCA 287

Db 3225 accccagcagggctctctgtccagaccctgcccgggcgctgggcatgtgtccacca 3284

Qy 288 CAGGCACCGCAGCTCATCTACCAGAGTGGGGTGGGACCTGACACTAGGGCTGGAGCC 347

Db 3285 caggcaccgcagctcatctaccagagtggtgggtggacctgacactagggctgagcc 3344

Qy 348 CTCTCAAGAGGAGGCCCGCCAGGCTCTCCAGTCCAGCCCTCCGAAAGGGCTGGCTCCGATGT 407

Db 3345 ctctgaagagaggccccccaggtctccactggcacctccgaaagggtggctccgagtc 3404

Qy 408 ATTTGATGTGACCTTGGGAATGGGGGCAGCCAAAGGGCTGCAAAAGCTCCCCACACATGA 467

Db 3405 atttgatgtgacctgggaatggggcagcaagggtgcaagacctccccacacatga 3464

Qy 468 CCCAGCCCTCTACAGCGGTACAGTAGAGACCCACAGTACCCCTGCCCTCTGTAGACTGA 527

Db 3465 cccagccctctacagcgtacagttagagacccacagtacccctgcccctctgagactga 3524

Qy 528 TGGCTACGTTGCCCCCTTCACCTGCAGCCCGCCAGCTGAATATGTGAACCCAGCAGATGT 587

Db 3525 tggctacgttgccccctgacctgagcccccagcgtgaatatgtgaacccagccagatgt 3584

Qy 588 TGGGCCCCAGCCCTTGCCTCCCGAGAGGGCCCTCTGCTGCTGCCCGACCTGCTGGTGC 647

Db 3585 tcggcccccagcccttcgcccagagggccctctgctgctgcccagcctgtgtgtgc 3644

Qy 648 CACTCTGGAAGGCCCAAGACTCTCTCCCGAGGGAAGTAATGGGGTCTCAAGACGCTTTT 707

Db 3645 cactctggaaggggccaaagactctctcccaggggaagaatgggtcgctcaagagcgtttt 3704

Qy 708 TGCCTTTGGGGTGCCTGGAGAACCCCGAG-ACCTGACACCCCGAGGAGGAGTGCCTCC 766

Db 3705 tgcctttgggggtgcgtggagaaaccccgagctctgacaccccccaggaggagcgcctcc 3764

Qy 767 TCAGCCCCACCTCTCTGCTTCAGCCCGACCTTCGACAACTCTATTACTGGGACCA 826

Db 3765 tcagcccccacccctctctgcttcagcccagccttcgacaacctctattactggacca 3824

Qy 827 GGACCCACAGAGCGGGGGTCCACCCAGACCTTCAAGGGACACCTACCGGCAGAGAA 886

Db 3825 ggacccaccagagcggggggtcccccagacacttcaaaaggagacactacggcagagaa 3884

Qy 887 CCCAGGTACCTGGGTCTGGAGCTGCCAGCTTGAAGCCCTTAAGGGCCCATATGCTGAGTGG 946

Db 3885 cccagagtacctgggtctggagcgtgccagtgtaaccagaaaggccaagtcgcgagaagcc 3944

Qy 947 ATGCTTGACCCACAGG 962

Db 3945 ctgatgtgtcctcagg 3960

RESULT 10
AAAL4812

ID XX AAA14812 standard; cDNA; 4472 BP.
 AC XX AAA14812;
 DT DT 08-AUG-2000 (first entry)
 DE XX cDNA encoding the SPLICE erbB-2 receptor protein.
 DE XX SPLICE erbB-2 receptor protein; cell transformation disorder; cancer;
 KW tumor cell proliferation; tissue degeneration; arthropathy;
 KW bone resorption; inflammatory disease; degenerative disorder;
 KW wound healing; ss.
 XX OS Homo sapiens.
 XX FH Key Location/Qualifiers
 FT CDS 175..3942
 FT /*tag= a
 FT /product= "SPLICE erbB-2 protein"
 XX PN WO200020579-A1.
 XX PD 13-APR-2000.
 XX PF 01-OCT-1999; 99WO-CA00912.
 XX PR 02-OCT-1998; 98US-0165192.
 XX XX (UTMC-) UNIV MCMASTER.
 XX PI Muller WJ, Siegel PM;
 XX DR WPI; 2000-303768/26.
 XX DR P-PSDB; AAY84780.
 XX XX Nucleic acid encoding an erbB 2 receptor protein designated SPLICE
 PT erbB-2, inhibitors of the protein are useful for treatment of cancer -
 XX Claim 4; Fig 1; 60pp; English.
 CC The present sequence encodes a SPLICE erbB-2 receptor protein. The
 CC protein has an in-frame deletion of 16 amino acids, 2 of which are
 CC conserved cysteine residues, compared to the unspliced protein. The
 CC erbB-2 polynucleotide is used to construct probes for detecting
 CC disorders of cell transformation such as cancer. Antibodies to the
 CC protein may be used to detect SPLICE erbB-2 in a sample. Agents
 CC (e.g. antisense oligonucleotides) which inhibit the expression of
 CC SPLICE erbB-2 are useful for reducing tumor cell proliferation and
 CC treating cancer. Substances which stimulate SPLICE erbB-2 are useful
 CC for treating conditions involving damaged cells including conditions
 CC in which degeneration of tissue occurs, such as arthropathy, bone
 CC resorption, inflammatory diseases, degenerative disorders of the
 CC central nervous system and wound healing.
 XX OS Sequence 4472 BP; 902 A; 1383 C; 1328 G; 859 T; 0 other;
 SQ
 Query Match 31.5%; Score 669; DB 21; Length 4472;
 Best Local Similarity 98.4%; Pred. NO. 6.5e-162;
 Matches 686; Conservative 0; Mismatches 10; Indels 1; Gaps 1;
 QY 228 ACGGACGACGAGATACATGTCTCCAGACCTGCGGGGCGGTGGGGACCTACACTAGGCGTGGAGCC 287
 Db 3249 accccagcagggtctctctctccagaccctgccggcgctgggggctggtccacca 3308
 QY 288 CAGGACCGGAGTCACTACAGGAGTGGCGGTGGGGACCTACACTAGGCGTGGAGCC 347
 Db 3309 caggcaccgagcttccatctaccaggagtgcgtgggagcctgcacactaggctggagcc 3368
 QY 348 CTCTGAGAGAGGAGCCGCCAGGCTCTCCAGTGGCACCTCCGAGGGGCTGGCTCCGATGT 407
 Db 3369 ctctgaagagaggccccaggtctctccactgcacccctccgaagggtggtccgatgt 3428

QY 408 ATTTGATGTTGACCTGGGAATGGGSCAGCCCAAGGGCTGCAAGCCCTCCCCACACATGA 467
 Db 3429 atttgatgtgacctgggaatggggcagcgaagggtgcaaaagcctccccacacatga 3488
 QY 468 CCCAGCCCTCTACAGCGGTACAGTGAGGACCCACACAGTACCCCTGCGCTCTGAGACTGA 527
 Db 3489 cccagccctctacagcggtacagtggagacccacagtagccctgcctctgagactga 3548
 QY 528 TGGCTACGTTGCGCCCTGACCTGCGAGCCCGCCAGCGCTGAATATGTGAACCGCAGATGT 587
 Db 3549 tggctacgttgccccctgacctgcagccccagcctgaatatgtgaaccagcagatgt 3608
 QY 588 TCGGCCCCAGGCCCTTCGCCCCGAGAGGGCCCTCTGCTGCTGCCGACCTGCTGGTGC 647
 Db 3609 tcggcccccagccctctgcgcccgagagggccctctgcctgcgcagacctgctggtgc 3668
 QY 648 CACTCTGGAAGGCCCAAGACTCTCTCCCGAGGAGTCTCTCCCGAGGAGAAATGGGGTCTGTC 707
 Db 3669 cactctggaaaggcccaagactctctcccgaggaatggggtctgcaaaagacgtttt 3728
 QY 708 TGCCTTTGGGGTCCCGTGGAGAACCCCGAG-ACCTTGACACCCCGAGGAGGAGTCCGCC 766
 Db 3729 tgcctttgggggtgcctggagaaaccccgagtacttgacaccccgagggagctgcccc 3788
 QY 767 TCAGCCCGACCTCTCTGCTGCTGAGCCAGCCTTCGACAACTCTATTACTGGGACCA 826
 Db 3789 tcagccccacctctctctccttcagccagccttcgacacacctctattactgggacca 3848
 QY 827 GGACCCACGAGCGGGGCTCCACCCAGCAGCTTCAAGGGGACACCTTACCGGACAGAA 886
 Db 3849 ggaccacacagagcgggggtccacccagaccttcaaggagacacacctacgagagaa 3908
 QY 887 CCCAGAGTACTTGGGTCTGGAGCTGCCAGTGTGAAGC 923
 Db 3909 ccagagtagctgggtggtgagctgcccagtgccagtggaacc 3945
 RESULT 11
 AAT40739
 ID AAT40739 standard; cDNA; 3768 BP.
 AC AAT40739;
 DT 01-JAN-1997 (first entry)
 XX HER-2/neu oncogene.
 XX DE HER-2/neu oncogene.
 KW HER-2/neu; c-erbB1; p185; oncogene; tyrosine protein kinase;
 KW breast cancer; ovary cancer; colon cancer; lung cancer;
 KW prostate cancer; genetic immunisation; tumour; vaccine; vector;
 KW ss.
 XX OS Homo sapiens.
 XX FH Key Location/Qualifiers
 FT CDS 1..3765
 FT /*tag= b
 FT /note= "nucleotides 2026-3765 (claim 1) code for
 HER-2/neu intracellular domain"
 XX PN WO9630514-A1.
 XX PD 03-OCT-1996.
 XX PF 28-MAR-1996; 96WO-US01689.
 XX PR 31-MAR-1995; 95US-0414417.
 XX XX (UNIV) UNIV WASHINGTON.
 XX PA Cheever MA, Disis ML;
 XX PI WPI; 1996-455361/45.
 DR

DR P-PSDB; AAW01111.

XX DNA encoding HER-2/neu poly:peptide(s) - used for prevention or

PT treatment of malignancies with which the HER-2/neu oncogene is

PT associated

XX Claim 1; Page 49-56; 71pp; English.

XX Human HER-2/neu oncogene cDNA (AAT40739) codes for HER-2/enu (p185 or

CC c-erbB2) protein (AAW01111). The oncogene is overexpressed in various

CC cancers, including breast, ovarian, colon, lung and prostate, and

CC appears to induce malignancies through quantitative mechanisms that

CC result from increased or deregulated expression of an essentially

CC normal gene product. Nucleotides 2026-3765 of the cDNA sequence

CC code for the intracellular domain (Lys676-Val1255) of the HER-2/neu

CC protein, which is useful for immunisation against malignancy.

CC Nucleic acids can be used to direct expression of the intracellular

CC domain in transformed host cells, or are used, alone or in a viral

CC vector, for genetic immunisation of an animal.

XX

SQ Sequence 3768 BP; 759 A; 1171 C; 1119 G; 719 T; 0 other;

Query Match 31.4%; Score 667.6; DB 17; Length 3768;

Best Local Similarity 98.6%; Pred. No. 1.4e-161;

Matches 684; Conservative 0; Mismatches 9; Indels 1; Gaps 1;

QY 228 ACGGACGAGAGATCAGATGTCAGAGCCCTGCCCGGCGCTGGGGCATGGTCCACCA 287

Db 3075 accccagcaggctctctctctcagaccctgccccggcgctggggcatggtccacca 3134

QY 288 CAGGACCGCAGCTCATCTACAGAGTGGCGGTGGGACCTGACACTAGGCTGGAGCC 347

Db 3135 caggcaccgagctcatctaccagagtgggcggtgggacctgacactagagctgagcc 3194

QY 348 CTCGTAAGAGAGGCCGCCAGGTCTCACTGGACCTCCGAGGGCTGGCTCGGATGT 407

Db 3195 ctctgaagagaggccccaggtctccactggcacctccgaagggtggctcgatgt 3254

QY 408 ATTTCATGCTGACCTGGGAATGGGGAGCAAGGGGTGCAAGCCCTCCACACATGA 467

Db 3255 atttgatggtgacctgggaatggggagcccaagggtgcaaaagctcccccacacatga 3314

QY 468 CCCAGCCCTCTACAGCGGTACAGTGGAGACCCACAGTACCTGCTGCTGCTGGTGC 527

Db 3315 cccagccctctacagcgtacagtgaagaccacagatccctgacctgtgagactga 3374

QY 528 TGGGTACGTTGCCGCCCTGAGCTGAGCCGCCAGCCTGAATATGTGAACCCAGAGATGT 587

Db 3375 tggctacgttgccccctgacctgagccccagcctgaatatgtgaaccagcagatgt 3434

QY 588 TCGGCCCCAGCCCTTCGCCCCGAGAGGGCCCTGCTGCTGCCCGACCTGGTGGTC 647

Db 3435 tcggccccagcccccttcgcccagagaggccccctgctgctgcccagacctgctgtgc 3494

QY 648 CACTGTGAAGAGGCCCAAGTCTCTCCAGGGAAGAAATGGGTGCTCAAGAGCTTTT 707

Db 3495 cactctgaaaggccccagactctctcccagggaagaatgggtctgcaaaagcgtttt 3554

QY 708 TGCCTTTGGGGTGGCGTGGAGAACCCCGAG-ACCTTGACCCCGAGGAGGAGTGC 766

Db 3555 tgcctttgggggtgctggagaaaccccgagtacttgacacccccaggagagctgcccc 3614

QY 767 TCAGGCCACCTCTCTCTGCTTCAGCCCGAGCCTTCACACACCTCTATTACTGGACCA 826

Db 3615 tcagccccacctctctctgcttcagccccagccttcgacaacctctattactgggacca 3674

QY 827 GGACCCACAGAGGGGGGGTCCACCCAGACCTTCAAGGGACACACTAGGCGAGAA 886

Db 3675 ggaccacacagagcgggggtccaccacagacacttcaaaaggacacctacggaagaa 3734

QY 887 CCCAGAGTACCTGGGTCTGGAGCTGCCAGTGTGA 920

Db 3135 caggcaccgagctcatctaccagagtgggcggtgggacctgacactagggctgagcc 3194

Db 3735 cccagagtacctgggtgctggacgtgccagtgta 3768

RESULT 12

AXX01912

ID AAX01912 standard; DNA; 3768 BP.

XX

AC AAX01912;

XX

DT 21-APR-1999 (first entry)

XX

DE Human HER-2/neu oncogene DNA.

XX

KW HER-2/neu; oncogene; immune response; T cell; B cell; immunisation;

KW malignancy; treatment; tumour; ss.

XX

OS Homo sapiens.

XX

FH Key Location/Qualifiers

FT CDS 1..3768

FT /*tag= a

FT /product= "HER-2/neu"

FT /note= "oncogene"

FT misc_feature 2026..3765

FT /*tag= b

FT /note= "region which elicits immune response"

XX

PN US5869445-A.

XX

PD 09-FEB-1999.

XX

PF 01-APR-1996; 96US-0625101.

XX

PR 01-APR-1996; 96US-0625101.

PR 17-MAR-1993; 93US-0033644.

PR 12-AUG-1993; 93US-0106112.

PR 31-MAR-1995; 95US-0414417.

XX

PA (UNIW) UNIV WASHINGTON.

XX

PI Cheever MA, Disis ML;

XX

DR WPI; 1999-152835/13.

DR P-PSDB; AAW92406.

XX

PT Use of HER-2/neu polypeptides - for eliciting an immune response to

PT an HER-2/neu associated malignancy, particularly for treating or

PT preventing tumours

XX

PS Claim 1a; Column 23-32; 26pp; English.

XX

CC This sequence encodes the human HER-2/neu oncogene protein. A fragment

CC of this protein is used in a method for eliciting or enhancing an immune

CC response to HER-2/neu protein. The polypeptide can stimulate T cells and

CC B cells to produce an immune response to the HER-2/neu protein. The

CC method can be used for immunisation against a malignancy in which the

CC HER-2/neu oncogene is associated and in the treatment of an existing

CC tumour, or to prevent tumour occurrence or reoccurrence.

XX

SQ Sequence 3768 BP; 759 A; 1171 C; 1119 G; 719 T; 0 other;

Query Match 31.4%; Score 667.6; DB 20; Length 3768;

Best Local Similarity 98.6%; Pred. No. 1.4e-161;

Matches 684; Conservative 0; Mismatches 9; Indels 1; Gaps 1;

QY 228 ACGGACGAGAGATCAGATGTCAGAGCCCTGCCCGGCGCTGGGGCATGGTCCACCA 287

Db 3075 accccagcaggctctctctcagaccctgccccggcgctggggcatggtccacca 3134

QY 288 CAGGACCGCAGCTCATCTACAGAGTGGCGGTGGGACCTGACACTAGGCTGGAGCC 347

Db 3135 caggcaccgagctcatctaccagagtgggcggtgggacctgacactagagctgagcc 3194

QY 348 CTCGTAAGAGAGGCCGCCAGGTCTCACTGGACCTCCGAGGGCTGGCTCGGATGT 407

Db 3195 ctctgaagagaggccccaggtctccactggcacctccgaagggtggctcgatgt 3254

QY 408 ATTTCATGCTGACCTGGGAATGGGGAGCAAGGGGTGCAAGCCCTCCACACATGA 467

Db 3255 atttgatggtgacctgggaatggggagcccaagggtgcaaaagctcccccacacatga 3314

QY 468 CCCAGCCCTCTACAGCGGTACAGTGGAGACCCACAGTACCTGCTGCTGCTGGTGC 527

Db 3315 cccagccctctacagcgtacagtgaagaccacagatccctgacctgtgagactga 3374

QY 528 TGGGTACGTTGCCGCCCTGAGCTGAGCCGCCAGCCTGAATATGTGAACCCAGAGATGT 587

Db 3375 tggctacgttgccccctgacctgagccccagcctgaatatgtgaaccagcagatgt 3434

QY 588 TCGGCCCCAGCCCTTCGCCCCGAGAGGGCCCTGCTGCTGCCCGACCTGGTGGTC 647

Db 3435 tcggccccagcccccttcgcccagagaggccccctgctgctgcccagacctgctgtgc 3494

QY 648 CACTGTGAAGAGGCCCAAGTCTCTCCAGGGAAGAAATGGGTGCTCAAGAGCTTTT 707

Db 3495 cactctgaaaggccccagactctctcccagggaagaatgggtctgcaaaagcgtttt 3554

QY 708 TGCCTTTGGGGTGGCGTGGAGAACCCCGAG-ACCTTGACCCCGAGGAGGAGTGC 766

Db 3555 tgcctttgggggtgctggagaaaccccgagtacttgacacccccaggagagctgcccc 3614

QY 767 TCAGGCCACCTCTCTCTGCTTCAGCCCGAGCCTTCACACACCTCTATTACTGGACCA 826

Db 3615 tcagccccacctctctctgcttcagccccagccttcgacaacctctattactgggacca 3674

QY 827 GGACCCACAGAGGGGGGGTCCACCCAGACCTTCAAGGGACACACTAGGCGAGAA 886

Db 3675 ggaccacacagagcgggggtccaccacagacacttcaaaaggacacctacggaagaa 3734

QY 887 CCCAGAGTACCTGGGTCTGGAGCTGCCAGTGTGA 920

Db 3135 caggcaccgagctcatctaccagagtgggcggtgggacctgacactagggctgagcc 3194

QY 288 CAGGACCCGACGCTCATCTACAGGAGTGGCGGTGGGACCTGACACTAGGCTGGAGCC 347
 Db 3309 CAGGACCCGACGCTCATCTACAGGAGTGGCGGTGGGACCTGACACTAGGCTGGAGCC 3368
 QY 348 CTCTGAAGAGGAGGCCCCCAGCTCTCCACTGGCACCTCCGAAGGGCTGGCTCCGATGT 407
 Db 3369 CTTCTGAAGAGGAGGCCCCCAGCTCTCCACTGGCACCTCCGAAGGGCTGGCTCCGATGT 3428
 QY 408 ATTTGATGTGACCTCTGGGAATGGGGAGCCAAAGGGCTGCAAGGCTCCCCACACATGA 467
 Db 3429 ATTTGATGTGACCTCTGGGAATGGGGAGCCAAAGGGCTGCAAGGCTCCCCACACATGA 3488
 QY 468 CCCAGCCCTCTACAGCGTACAGTGAAGACCCACAGTACCCCTGCTGAGACTGA 527
 Db 3489 CCCAGCCCTCTACAGCGTACAGTGAAGACCCACAGTACCCCTGCTGAGACTGA 3548
 QY 528 TGGCTACGTTGCCCCCTGACCTGAGCCGCCAGCCCTGAATATGTGAACAGCCAGATGT 587
 Db 3549 TGGCTACGTTGCCCCCTGACCTGAGCCGCCAGCCCTGAATATGTGAACAGCCAGATGT 3608
 QY 588 TGGGCCCCAGCCCCCTTGGCCCCGAGAGGGCCCTCTGCTGCTGCCCCGACCTGCTGGTGC 647
 Db 3609 TGGGCCCCAGCCCCCTTGGCCCCGAGAGGGCCCTCTGCTGCTGCCCCGACCTGCTGGTGC 3668
 QY 648 CACTCTGGAAGGCCCCCAAGACTCTCTCCAGGAGGAGATGGGTGCTCAAGAGCTTTT 707
 Db 3669 CACTCTGGAAGGCCCCCAAGACTCTCTCCAGGAGGAGATGGGTGCTCAAGAGCTTTT 3728
 QY 708 TGCCCTTTGGGGTGGCGTGGAGAACCCCGAG-ACVTGACACCCAGGAGGAGCTGCCCC 766
 Db 3729 TGCCCTTTGGGGTGGCGTGGAGAACCCCGAGTACTTACACCCAGGAGGAGCTGCCCC 3788
 QY 767 TCAGCCCCACCTCTCTGCTTGCCTTCAGCCCCAGCCCTTCGACACCTCTATTACTGGGACCA 826
 Db 3789 TCAGCCCCACCTCTCTGCTTGCCTTCAGCCCCAGCCCTTCGACACCTCTATTACTGGGACCA 3848
 QY 827 GGACCCACAGAGCGGGGGTCTCCAGCCAGACCTTCAAGAGGACACCTAGGCGAGAGAA 886
 Db 3849 GGACCCACAGAGCGGGGGTCTCCAGCCAGACCTTCAAGAGGACACCTAGGCGAGAGAA 3908
 QY 887 CCCAGTACCTGGGTCTGGAGTCCGACGCTTGAAGCCCTTAAGGGCCATATGGTGAGTGG 946
 Db 3909 CCCAGTACCTGGGTCTGGAGTCCGACGCTTGAAGCCCTTAAGGGCCATATGGTGAGTGG 3968
 QY 947 ATGCTTGACCCCGAGG 962
 Db 3969 CTGATGTGCTCTCAGG 3984

RESULT 2

US-09-056-105-26
 ; Sequence 26, Application US/09056105
 ; Patent No. 6287569
 ; GENERAL INFORMATION:
 ; APPLICANT: KIPPS, THOMAS J.
 ; APPLICANT: WU, YUNQI
 ; TITLE OF INVENTION: VACCINES WITH ENHANCED INTRACELLULAR
 ; FILE REFERENCE: 233/221
 ; CURRENT APPLICATION NUMBER: US/09/056.105
 ; EARLIER FILING DATE: 1998-04-06
 ; EARLIER APPLICATION NUMBER: 60/043,467
 ; EARLIER FILING DATE: 1997-04-10
 ; NUMBER OF SEQ ID NOS: 35
 ; SOFTWARE: FastSeq for Windows Version 3.0
 ; SEQ ID NO 26
 ; LENGTH: 4473
 ; TYPE: DNA
 ; ORGANISM: Homo sapiens
 ; US-09-056-105-26

Query Match 31.6%; Score 671.2; DB 4; Length 4473;

Best Local Similarity 95.4%; Pred. No. 7.4e-175;
 Matches 702; Conservative 0; Mismatches 33; Indels 1; Gaps 1;
 QY 228 ACGGCAGCAGAGAATACATGTCCAGACCCCTGCCCGGCGCTGGGCGCATGGTCCACCA 287
 Db 3249 accccagcagggctctctctccagagacctgccccggcgctggggcatggtccaccac 3308
 QY 288 CAGGACCCGACGCTCATCTACAGGAGTGGGGTGGGACCTGACACTAGGCTGGAGCC 347
 Db 3309 caggcaccgagctcatctaccagagtgggcggtgggagacctgacctagggctggagcc 3368
 QY 348 CTCGTAAGAGAGGCCCCCAGGCTCTCCACTGGCACCTCCGAAGGGCTGGCTCCGATGT 407
 Db 3369 ctctgaagagagggccccccaggtctccactggcacctccgaagggctggctccgatgt 3428
 QY 408 ATTTGATGTGACCTCTGGGAATGGGGAGCCAAAGGGCTGCAAGGCTCCCCACACATGA 467
 Db 3429 atttgaatggagacctgggaatgggggagccaaagggtgcaaaagctccccacacatga 3488
 QY 468 CCCAGCCCTCTACAGCGTACAGTGAAGACCCACAGTACCCCTGCTGAGACTGA 527
 Db 3489 cccagccctctacagcgggtacagtgagagacccacagtagccctgacctgagactga 3548
 QY 528 TGGCTACGTTGCCCCCTGACCTGAGCCGCCAGCCCTGAATATGTGAACAGCCAGATGT 587
 Db 3549 tggctacgttgccccctgaacctgagccccagcctgaatatgtgaaccagccagatgt 3608
 QY 588 TCGGCCCCAGCCCCCTTGGCCCCGAGAGGGCCCTCTGCTGCTGCCGACCTGCTGGTGC 647
 Db 3609 tcggccccagcccccttcgccccgagagggccctctgctgctgccccgacctgctggcgc 3668
 QY 648 CACTCTGGAAGGCCCCCAAGACTCTCTCCAGGAGGAGATGGGTGCTCAAGAGCTTTT 707
 Db 3669 cactctggaaaggcccccaagactctctcccaggaagaatgggtctcaaaagacgtttt 3728
 QY 708 TGCCCTTTGGGGTGGCGTGGAGAACCCCGAG-ACVTGACACCCAGGAGGAGCTGCCCC 766
 Db 3729 tgccctttgggggtggcgaggaaaccccgagtacttgacacccagggagagctgcccc 3788
 QY 767 TCAGCCCCACCTCTCTGCTTGCCTTCAGCCCCAGCCCTTCGACACCTCTATTACTGGGACCA 826
 Db 3789 tcagccccacctctctctgcttgccttcagccccagccttcgaaacctctattactgggacca 3848
 QY 827 GGACCCACAGAGCGGGGGTCTCCAGCCAGACCTTCAAGAGGACACCTAGGCGAGAGAA 886
 Db 3849 ggacccaccagagcgggggctccacccagcacctcaaaaggagacacctacgagagaa 3908
 QY 887 CCCAGTACCTGGGTCTGGAGTCCGACGCTTGAAGCCCTTAAGGGCCATATGGTGAGTGG 946
 Db 3909 cccagagtacctgggtctggagctgcccagtgcccagtggaaccagagggccaagtccgcagaagcc 3968
 QY 947 ATGCTTGACCCCGAGG 962
 Db 3969 ctgatgtgctctcagg 3984

RESULT 3
 US-08-229-515a-9
 ; Sequence 9, Application US/08229515a
 ; Patent No. 5518885
 ; GENERAL INFORMATION:
 ; APPLICANT: RAZIUDIN
 ; APPLICANT: SARKAR, FAZLUL H
 ; TITLE OF INVENTION: ERBB2 PROMOTER BINDING PROTEIN IN
 ; TITLE OF INVENTION: NEOPLASTIC DISEASE
 ; NUMBER OF SEQUENCES: 19
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: NEEDLE & ROSENBERG PC
 ; STREET: 127 Peachtree Street, Suite 1200
 ; CITY: Atlanta
 ; STATE: Georgia
 ; COUNTRY: usa
 ; ZIP: 30303

```

; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/229,515A
; FILING DATE: 19 APR 1994
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: PERRYMAN, DAVID G
; REGISTRATION NUMBER: 33,438
; REFERENCE/DOCKET NUMBER: 1414.608
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 404-688-0770
; TELEFAX: 404-688-9880
; INFORMATION FOR SEQ ID NO: 9:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 4530 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
;
; US-08-229-515A-9

```

Query Match 31.5%; Score 669.6; DB 1; Length 4530;
Best Local Similarity 95.2%; Pred. No. 2.1e-174;
Matches 701; Conservative 0; Mismatches 34; Indels 1;
Gaps 1;

QY	228	ACGGCAGCAGAGATCACA	TTCCAGACCTTGC	CCCCGGGCCTGGGGCAT	TGGTCCACCA	287
Db	3225	ACCCACAGGGCTTTCT	TCTGTCCAGACCTTGC	CCCCGGGCCTGGGGCAT	TGGTCCACCA	3284
QY	288	CAGGCACCCAGCTCAT	CTTACCAGAGTGGCGT	GGGACCTGACACT	TAGGGCTGGAGCC	347
Db	3285	CAGGCACCCAGCTCAT	CTTACCAGAGTGGCGT	GGGACCTGACACT	TAGGGCTGGAGCC	3344
QY	348	CTCTGAAGAGAGGCCCC	AGGCTTCCACTGGCAC	CCCTTCCGAAGGGCT	TGGCTCCGATGT	407
Db	3345	CTCTGAAGAGAGGCCCC	AGGCTTCCACTGGCAC	CCCTTCCGAAGGGCT	TGGCTCCGATGT	3404
QY	408	ATTTGATGTGTGACCT	TGGGAATGGGGGAC	CCAAAGGGCTCAAA	AGCCTCCCCACACATGA	467
Db	3405	ATTTGATGTGTGACCT	TGGGAATGGGGGAC	CCAAAGGGCTCAAA	AGCCTCCCCACACATGA	3464
QY	468	CCCCAGCCCTCTACAG	CGGTACAGTGAGGAC	CCCCACAGTACCCCT	TGCCCTCTTGAGACTGA	527
Db	3465	CCCCAGCCCTCTACAG	CGGTACAGTGAGGAC	CCCCACAGTACCCCT	TGCCCTCTTGAGACTGA	3524
QY	528	TGGCTACGTTGGCCCC	CTGACCTCAGCCCC	CAGCCTGAATATGT	GAACAGCCAGATGT	587
Db	3525	TGGCTACGTTGGCCCC	CTGACCTCAGCCCC	CAGCCTGAATATGT	GAACAGCCAGATGT	3584
QY	588	TGGCCCCCAGCCCCCT	TGCCCCGAGAGGGCC	CTTGCTCTGCCCGAC	CTCTGCTGCTGC	647
Db	3585	TGGCCCCCAGCCCCCT	TGCCCCGAGAGGGCC	CTTGCTCTGCCCGAC	CTCTGCTGCTGC	3644
QY	648	CACCTCGAAAGGCCCA	AGACTCTCTCCCC	CAGGAGAAATGGGGT	TCGTCAAAGACGTTTT	707
Db	3645	CACCTCGAAAGGCCCA	AGACTCTCTCCCC	CAGGAGAAATGGGGT	TCGTCAAAGACGTTTT	3704
QY	708	TGGCTTTGGGGTGCGG	TGGAGAACCCCGAG	-ACTTGGACCC	CAGGAGGAGCTGCCCC	766
Db	3705	TGGCTTTGGGGTGCGG	TGGAGAACCCCGAG	TACTTGACCC	CAGGAGGAGCTGCCCC	3764
QY	767	TCAGCCCCACCTCTCT	CTTGCCCTTCAGCC	CCAGCCTTCGACAA	CCCTCTATTACTGGGACCA	826
Db	3765	TCAGCCCCACCTCTCT	CTTGCCCTTCAGCC	CCAGCCTTCGACAA	CCCTCTATTACTGGGACCA	3824
QY	827	GGACCCACAGAGGGGG	GGCTTCCACCC	AGCACCTTCAAG	AGGGACCTACGGCAGAGAA	886
Db	3825	GGACCCACAGAGGGGG	GGCTTCCACCC	AGCACCTTCAAG	AGGGACCTACGGCAGAGAA	3884

```

QY      887   CCACAGTAGTACTGGGTCTCGGACGTGTGAGCCCTTAAGGGCCCATATGGTGAGTGG    946
Db      3885   CCCAGATTACTTGGGTCTCGGAGCTGCCAGTGTGAACCACGAGAGGCCAAGTCGCCAGAAGCC    3944
QY      947   ATGCCTTGACCCAGG    962
Db      3945   CTGATGTGCTCCTCAGG    3960

RESULT          4
US-08-645-865-9
; Sequence 9, Application US/08645865
; Patent No. 5654406
; GENERAL INFORMATION:
; APPLICANT: RAZIUDIN
; APPLICANT: SARKAR, FAZLUL H
; TITLE OF INVENTION: ERBB2 PROMOTER BINDING PROTEIN IN
; TITLE OF INVENTION: NEOPLASTIC DISEASE
; NUMBER OF SEQUENCES: 19
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: NEEDLE & ROSENBERG PC
; STREET: 127 Peachtree Street, Suite 1200
; CITY: Atlanta
; STATE: Georgia
; COUNTRY: usa
; ZIP: 30303
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/645,865
; FILING DATE: 14 MAY 1996
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: PERRYMAN, DAVID G
; REGISTRATION NUMBER: 33,438
; REFERENCE/DOCKET NUMBER: 1414.608
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 404-688-0770
; TELEFAX: 404-688-9880
; INFORMATION FOR SEQ ID NO.: 9:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 4530 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; US-08-645-865-9

```

Query Match	31.5%	Score	669.6;	DB 1;	Length	4530;			
Best Local Similarity	95.2%;	Presd.	No. 2.1e-174;						
Matches	701;	Conservative	0;	Mismatches	34;	Indels	1;	Gaps	1;
QY	228	ACGGCAGCAGAGATCATCTTCAGACCCCTGCCCGGGCGCTGGGGCGCATGGTCCACCA	287						
Db	3225	ACCCACGAGGGCTTCCTCTTCAGACCCCTGCCCGGGCGCTGGGGCGCATGGTCCACCA	3284						
QY	288	CAGGCACCGCAGCTCATCTACACGAGTGGCGGTGGGACCTGCACCTPAGGGCTGGAGCC	347						
Db	3285	CAGGCACCGCAGCTCATCTACCAGAGTGGCGGTGGGACCTGCACCTAGGCTGGAGCC	3344						
QY	348	CTCTGAAGAGAGGCCGCCAGGTTCTCACTGGCACCTCTCCGAAGGGGTGGCTCGGATGT	407						
Db	3345	CTCTGAAGAGAGGCCGCCAGGTTCTCACTGGCACCTCTCCGAAGGGGTGGCTCGGATGT	3404						
QY	408	ATTGTATGGTTCACCTGGGAATGGGGGAGCAAGGGGTGCAAGCCCTCCCCACACATGA	467						
Db	3405	ATTGTATGGTGACCTGGGAATGGGGGAGCAAGGGGTGCAAGCCCTCCCCACACATGA	3464						
QY	468	CCCAGGCCCTCTACAGGGGTACGTGAGGACCCACACAGTACCCCTGCCCTCTGAGACTGA	527						

NUMBER OF SEQUENCES: 16
CORRESPONDENCE ADDRESS:
ADDRESSEE: Edmund R. Pitcher, Testa, Hurwitz, & Thibault
STREET: Exchange Place, 53 State Street
CITY: Boston
STATE: Massachusetts
COUNTRY: USA
ZIP: 02109

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/356,786
FILING DATE:

CLASSIFICATION: 424

PRIOR APPLICATION DATA:

APPLICATION NUMBER: 07/831,967

FILING DATE: 06-FEB-1992

ATTORNEY/AGENT INFORMATION:

NAME: Pitcher, Edmund R.

REGISTRATION NUMBER: 27,829

REFERENCE/DOCKET NUMBER: CRP-053

TELECOMMUNICATION INFORMATION:

TELEPHONE: (617) 248-7000

TELEFAX: (617) 248-7100

INFORMATION FOR SEQ ID NO: 1:

SEQUENCE CHARACTERISTICS:

LENGTH: 3768 base pairs

TYPE: nucleic acid

STRANDEDNESS: single

TOPOLOGY: linear

MOLECULE TYPE: cDNA

FEATURE:

NAME/KEY: CDS

LOCATION: 1..3768

OTHER INFORMATION: /note= "product = "cerB-b2""

US-08-356-786-1

Query Match 31.4%; Score 667.6; DB 2; Length 3768;
Best Local Similarity 98.6%; Pred. No. 6.7e-174;
Matches 684; Conservative 0; Mismatches 9; Indels 1; Gaps 1;

QY 228 ACGGACGAGAGATCATCTTCCAGACCTGCCCGGGCGCTGGGGCATGGTCCACCA 287
DB 3075 ACCCCAGCAGGGCTTCTTGTCCAGACCTGCCCGGGCGCTGGGGCATGGTCCACCA 3134
QY 288 CAGGCACCGCAGCTCATCTTACCAGGAGTGGCGGTGGGACCTGCACACTAGGGCTGGAGCC 347
DB 3135 CAGGCACCGCAGCTCATCTTACCAGGAGTGGCGGTGGGACCTGCACACTAGGGCTGGAGCC 3194
QY 348 CTCTGAAGAGGAGGCCCCAGGTCTTCCACTGCGACCCCTCCGAAGGGGCTGGCTCCGATGT 407
DB 3195 CTCTGAAGAGGAGGCCCCAGGTCTTCCACTGCGACCCCTCCGAAGGGGCTGGCTCCGATGT 3254
QY 408 ATTTGATGGTGACCTGGGAATGGGGGAGCCAAAGGGGTGCAAGCCCTCCACACATGA 467
DB 3255 ATTTGATGGTGACCTGGGAATGGGGGAGCCAAAGGGGTGCAAGCCCTCCACACATGA 3314
QY 468 CCCCAGCCCTCTACAGCGGTACAGTGAGGAGCCACAGTACCCCTGCCCTCTGAGACTGA 527
DB 3315 CCCCAGCCCTCTACAGCGGTACAGTGAGGAGCCACAGTACCCCTGCCCTCTGAGACTGA 3374
QY 528 TGGCTACGTGGCCCCCTGACCTGACGAGCCCCCAGCGCTGAATATGTGAACCCAGCCAGATGT 587
DB 3375 TGGCTACGTGGCCCCCTGACCTGACGAGCCCCCAGCGCTGAATATGTGAACCCAGCCAGATGT 3434
QY 588 TCGGCCCCAGCCCCCTTGGCCCCAGAGAGGGCCCTCTGCTGTGCTGCCCGGACCTGTGGTGC 647
DB 3435 TCGGCCCCAGCCCCCTTGGCCCCAGAGAGGGCCCTCTGCTGTGCTGCCCGGACCTGTGGTGC 3494

QY 648 CACTCTGGAAGGCCCAAGACTCTCTCCCGAGGAAGATGGGTCGTCAAAGAGCTTTT 707
DB 3495 CACTCTGGAAGGCCCAAGACTCTCTCCCGAGGAAGATGGGTCGTCAAAGAGCTTTT 3554
QY 708 TCCCTTTGGGGGTGCCGTGGAGAACCCCGAG-ACITGACACCCCGAGGAGGAGCTGCCCC 766
DB 3555 TGCCTTTGGGGGTGCCGTGGAGAACCCCGAGTACTTGACACCCCGAGGAGGAGCTGCCCC 3614
QY 767 TCAGCCGCCACCCCTCTCTTCCCTTCCAGCCCGAGCTTCGACAACCTCTATTACTGGGACCA 826
DB 3615 TCAGCCGCCACCCCTCTCTTCCCTTCCAGCCCGAGCTTCGACAACCTCTATTACTGGGACCA 3674
QY 827 GGACCCACCCAGAGCGGGGGCTCCACCCAGCACCTTCAAAGGGGACACCTTACGCGCAGAGAA 886
DB 3675 GGACCCACCCAGAGCGGGGGCTCCACCCAGCACCTTCAAAGGGGACACCTTACGCGCAGAGAA 3734
QY 887 CCCAGAGTACCTGGGTCTGGACGTGCCAGTGTGA 920
DB 3735 CCCAGAGTACCTGGGTCTGGACGTGCCAGTGTGA 3768

RESULT 7

US-08-229-515A-14

Sequence 14, Application US/08229515A

Patent No. 5518885

GENERAL INFORMATION:

APPLICANT: RAZIUDIN

APPLICANT: SARKAR, FAZLUL H

TITLE OF INVENTION: ERB2 PROMOTER BINDING PROTEIN IN

TITLE OF INVENTION: NEOPlastic DISEASE

NUMBER OF SEQUENCES: 19

CORRESPONDENCE ADDRESS:

ADDRESSEE: NEEDLE & ROSENBERG PC

STREET: 127 Peachtree Street, Suite 1200

CITY: Atlanta

STATE: Georgia

COUNTRY: USA

ZIP: 30303

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: PatentIn Release #1.0, Version #1.30

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/229,515A

FILING DATE: 19 APR 1994

CLASSIFICATION: 435

ATTORNEY/AGENT INFORMATION:

NAME: PERRYMAN, DAVID G

REGISTRATION NUMBER: 33,438

REFERENCE/DOCKET NUMBER: 1414.608

TELECOMMUNICATION INFORMATION:

TELEPHONE: 404-688-0770

TELEFAX: 404-688-9880

INFORMATION FOR SEQ ID NO: 14:

SEQUENCE CHARACTERISTICS:

LENGTH: 3955 base pairs

TYPE: nucleic acid

STRANDEDNESS: single

TOPOLOGY: linear

US-08-229-515A-14

Query Match 23.3%; Score 495.8; DB 1; Length 3955;

Best Local Similarity 83.0%; Pred. No. 1.2e-126;

Matches 577; Conservative 0; Mismatches 117; Indels 1; Gaps 1;

QY 232 CAGCAGACATCATGTCCAGACCTGCCCGGGCGCTGGGGCATGGTCCACCCACAGG 291

DB 3110 CAGCAGGAGTATCTTCTCCCGGACCTTACCCAGGACCTGGGAGCACACCCATAGAAGG 3169

QY 292 CACCAGCTCATCTACCAAGGAGTGGCGGTGGGGACCTGCACACTAGGGCTGGAGCCCTCT 351


```

; TITLE OF INVENTION: HUMAN NEUTRALIZING MONOCLONAL ANTIBODIES
; TITLE OF INVENTION: TO HUMAN IMMUNODEFICIENCY VIRUS
; NUMBER OF SEQUENCES: 170
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: The Scripps Research Institute, Office of
; ADDRESSEE: Patent Counsel
; STREET: 10666 No. 5652138th Torrey Pines Road, Suite 220,
; CITY: La Jolla
; STATE: CA
; COUNTRY: USA
; ZIP: 92037
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/276,852
; FILING DATE: 18-JUL-1994
; CLASSIFICATION: 514
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/178,302
; FILING DATE: 30-SEP-1993
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/954,148
; FILING DATE: 30-SEP-1992
; ATTORNEY/AGENT INFORMATION:
; NAME: Fitting, Thomas
; REGISTRATION NUMBER: 34,163
; REFERENCE/DOCKET NUMBER: SCR1452P
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 619-554-2937
; TELEFAX: 619-554-6312
; INFORMATION FOR SEQ ID NO: 156:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 13254 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: circular
; MOLECULE TYPE: DNA (genomic)
; US-08-276-852-156

```

```

Query Match      21.4%; Score 454.8; DB 1; Length 13254;
Best Local Similarity 99.6%; Pred. No. 4e-115;
Matches 456; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1411 GTGTAGCCGTAGTTAGGCCACCACTTCAAGAACTCTGTAGCAGCGCTACATACCTCGCT 1470
DB 5449 GTGTAGCCGTAGTTAGGCCACCACTTCAAGAACTCTGTAGCAGCGCTACATACCTCGCT 6390
QY 1471 CTGCTAATCCTGTTACCACTGCTGTCGCCAGTGGCGATAAGTCGTCTTACCGGGTTG 1530
DB 6389 CTGCTAATCCTGTTACCACTGCTGTCGCCAGTGGCGATAAGTCGTCTTACCGGGTTG 6330
QY 1531 GACTCAAGACGATAGTTACCGGATAAGGCGCAGCGCTGCGGCTGAACGGGGGTTGCTGC 1590
DB 6329 GACTCAAGACGATAGTTACCGGATAAGGCGCAGCGCTGCGGCTGAACGGGGGTTGCTGC 6270
QY 1591 ACACAGCCCGCTGTTAGGAGCAACGACCTACCGAACTGAGATACCTACAGCGGTGAGCAT 1650
DB 6269 ACACAGCCCGCTGTTAGGAGCAACGACCTACCGAACTGAGATACCTACAGCGGTGAGCAT 6210
QY 1651 TGAGAAAGCGCCACGCTTCCCGAAGGAGGAGAAAGCGGACAGGTATCCGGTAAAGCGGCAGG 1710
DB 6209 TGAGAAAGCGCCACGCTTCCCGAAGGAGGAGAAAGCGGACAGGTATCCGGTAAAGCGGCAGG 6150
QY 1711 CTCGGNACAGGAGAGCGCACGAGGAGCTTCCAGGGGGAAGCCCTGATCTTTATAGT 1770
DB 6149 CTCGGNACAGGAGAGCGCACGAGGAGCTTCCAGGGGGAAGCCCTGATCTTTATAGT 6090
QY 1771 CCTGTGCGGGTTTCGCCACCTCTCTGTAGCGCTGCGGCTGCGGCTGCGTCAAGGGGGG 1830

```

```

DB 6089 CCTGTGCGGGTTTCGCCACCTCTGACTTGAGCGTCGATTTTGTGATGCTCGTCAGGGGG 6030
QY 1831 CGGAGCCTATGGAAGAACGCCAGCAACGCCGCGCGGGG 1868
DB 6029 CGGAGCCTATGGAAGAACGCCAGCAACGCCGCGCGGGG 5992

RESULT 10
US-08-276-852-170
; Sequence 170, Application US/08276852
; Patent No. 5652138
; GENERAL INFORMATION:
; APPLICANT: Burton, Dennis R
; APPLICANT: Barbas, Carlos F
; APPLICANT: Lerner, Richard A
; TITLE OF INVENTION: HUMAN NEUTRALIZING MONOCLONAL ANTIBODIES
; TITLE OF INVENTION: TO HUMAN IMMUNODEFICIENCY VIRUS
; NUMBER OF SEQUENCES: 170
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: The Scripps Research Institute, Office of
; ADDRESSEE: Patent Counsel
; STREET: 10666 No. 5652138th Torrey Pines Road, Suite 220,
; CITY: La Jolla
; STATE: CA
; COUNTRY: USA
; ZIP: 92037
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/276,852
; FILING DATE: 18-JUL-1994
; CLASSIFICATION: 514
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/178,302
; FILING DATE: 30-SEP-1993
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/954,148
; FILING DATE: 30-SEP-1992
; ATTORNEY/AGENT INFORMATION:
; NAME: Fitting, Thomas
; REGISTRATION NUMBER: 34,163
; REFERENCE/DOCKET NUMBER: SCR1452P
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 619-554-2937
; TELEFAX: 619-554-6312
; INFORMATION FOR SEQ ID NO: 170:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 13254 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: circular
; MOLECULE TYPE: DNA (genomic)
; US-08-276-852-170

```

```

Query Match      21.4%; Score 454.8; DB 1; Length 13254;
Best Local Similarity 99.6%; Pred. No. 4e-115;
Matches 456; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1411 GTGTAGCCGTAGTTAGGCCACCACTTCAAGAACTCTGTAGCAGCGCTACATACCTCGCT 1470
DB 6806 GTGTAGCCGTAGTTAGGCCACCACTTCAAGAACTCTGTAGCAGCGCTACATACCTCGCT 6865
QY 1471 CTGCTAATCCTGTTACCACTGCTGTCGCCAGTGGCGATAAGTCGTCTTACCGGGTTG 1530
DB 6866 CTGCTAATCCTGTTACCACTGCTGTCGCCAGTGGCGATAAGTCGTCTTACCGGGTTG 6925
QY 1531 GACTCAAGACGATAGTTACCGGATAAGGCGCAGCGCTGCGGCTGAACGGGGGTTGCTGC 1590

```

```

Db 6926 GACTCAAGACGATAGTTACCGGATAAGCGCAGCGGTGCGGCTGAACGGGGGTTTCGTGC 6985
Qy 1591 ACACAGCCAGCTTGGAGCGAAGCACTTACACCGAAGTACAGTACCTACACGCTGAGCAT 1650
Db 6986 ACACAGCCAGCTTGGAGCGAAGCACTTACACCGAAGTACAGTACCTACACGCTGAGCAT 7045
Qy 1651 TGAGAAAGCGGACGCTTCCGAGAGGGAGAAAGCGGACAGGTATCCGGTAAGCGGAGG 1710
Db 7046 TGAGAAAGCGGACGCTTCCGAGAGGGAGAAAGCGGACAGGTATCCGGTAAGCGGAGG 7105
Qy 1711 GTCGGAACAGGAGCGCAGCAGAGGAGCTTCCAGGGGGAACGCTCGGTATCTTTATAGT 1770
Db 7106 GTCGGAACAGGAGCGCAGCAGAGGAGCTTCCAGGGGGAACGCTCGGTATCTTTATAGT 1765
Qy 1771 CTTGTCGGGTTTCGCCACCTCTGACTTGAGCGTGCATTTTGTGATGCTCGTCAGGGGGG 1830
Db 7166 CTTGTCGGGTTTCGCCACCTCTGACTTGAGCGTGCATTTTGTGATGCTCGTCAGGGGGG 1868
Qy 1831 CGGAGCCTATGGAAGAAAGCGCAGCAACGCGCGCGGGG 1868
Db 7226 CGGAGCCTATGGAAGAAAGCGCAGCAACGCGCGCGGGG 7263

```

RESULT 11

```

US-08-899-575-156/c
; Sequence 156, Application US/08899575
; Patent No. 5770440
; GENERAL INFORMATION:
; APPLICANT: Burton, Dennis R
; APPLICANT: Barbas, Carlos F
; APPLICANT: Lerner, Richard A
; TITLE OF INVENTION: HUMAN NEUTRALIZING MONOCLONAL ANTIBODIES
; TITLE OF INVENTION: TO HUMAN IMMUNODEFICIENCY VIRUS
; NUMBER OF SEQUENCES: 170
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: The Scripps Research Institute, Office of
; STREET: 10666 No. 5770440th Torrey Pines Road, Suite 220,
; CITY: La Jolla
; STATE: CA
; COUNTRY: USA
; ZIP: 92037
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/899,575
; FILING DATE: 18-JUL-1994
; APPLICATION NUMBER: US 08/178,302
; FILING DATE: 30-SEP-1993
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/954,148
; FILING DATE: 30-SEP-1992
; ATTORNEY/AGENT INFORMATION:
; NAME: Fitting, Thomas
; REGISTRATION NUMBER: 34,163
; REFERENCE/DOCKET NUMBER: SCR1452P
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 619-554-2937
; TELEFAX: 619-554-6312
; INFORMATION FOR SEQ ID NO: 156:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 13254 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double

```

```

; TOPOLOGY: circular
; MOLECULE TYPE: DNA (genomic)
; US-08-899-575-156
Query Match 21.4%; Score 454.8; DB 1; Length 13254;
Best Local Similarity 99.6%; Pred. No. 4e-115;
Matches 456; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
Qy 1411 GTGTAGCCCTAGTTAGGCCACCACTTCAAGAACTCTGTAGCACCCTACATACCTCGCT 1470
Db 6449 GTGTAGCCCTAGTTAGGCCACCACTTCAAGAACTCTGTAGCACCCTACATACCTCGCT 6390
Qy 1471 CTGCTAATCTCTTACCAGTGGCTGCTGCCACTGCGGTAAGTCTGTCTTACCGGGTTG 1530
Db 6389 CTGCTAATCTCTTACCAGTGGCTGCTGCCACTGCGGTAAGTCTGTCTTACCGGGTTG 6330
Qy 1531 GACTCAAGACGATAGTTACCGGATAAGCGCAGCGTCCGGCTGAACGGGGGTTTCGTGC 1590
Db 6329 GACTCAAGACGATAGTTACCGGATAAGCGCAGCGTCCGGCTGAACGGGGGTTTCGTGC 6270
Qy 1591 ACACAGCCAGCTTGGAGCGAAGCACTTACACCGAAGTACACCGAAGTACACCGTGAAGCAT 1650
Db 6269 ACACAGCCAGCTTGGAGCGAAGCACTTACACCGAAGTACACCGAAGTACACCGTGAAGCAT 6210
Qy 1651 TGAGAAAGCGCCACGCTTCCCGAAGGGAGAAAGCGGCGAGGTATCCCGTAAGCGGAGG 1710
Db 6209 TGAGAAAGCGCCACGCTTCCCGAAGGGAGAAAGCGGCGAGGTATCCCGTAAGCGGAGG 6150
Qy 1711 GTCGGAACAGGAGCGCAGCAGGAGGCTTCCAGGGGGAACGCTGTTATCTTTATAGT 1770
Db 6149 GTCGGAACAGGAGCGCAGCAGGAGGCTTCCAGGGGGAACGCTGTTATCTTTATAGT 6090
Qy 1771 CTTGTCGGGTTTCGCCACCTCTGACTTGAGCGTGCATTTTGTGATGCTCGTCAGGGGGG 1830
Db 6089 CTTGTCGGGTTTCGCCACCTCTGACTTGAGCGTGCATTTTGTGATGCTCGTCAGGGGGG 6030
Qy 1831 CGGAGCCTATGGAAGAAAGCGCAGCAACGCGCGCGGGG 1868
Db 6029 CGGAGCCTATGGAAGAAAGCGCAGCAACGCGCGCGGAGG 5992

```

RESULT 12

```

US-08-899-575-170
; Sequence 170, Application US/08899575
; Patent No. 5770440
; GENERAL INFORMATION:
; APPLICANT: Burton, Dennis R
; APPLICANT: Barbas, Carlos F
; APPLICANT: Lerner, Richard A
; TITLE OF INVENTION: HUMAN NEUTRALIZING MONOCLONAL ANTIBODIES
; TITLE OF INVENTION: TO HUMAN IMMUNODEFICIENCY VIRUS
; NUMBER OF SEQUENCES: 170
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: The Scripps Research Institute, Office of
; ADDRESSEE: Patent Counsel
; STREET: 10666 No. 5770440th Torrey Pines Road, Suite 220,
; CITY: La Jolla
; STATE: CA
; COUNTRY: USA
; ZIP: 92037
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/899,575
; FILING DATE: 24-JUL-1997
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/276,852

```

; FILING DATE: 18-JUL-1994
; APPLICATION NUMBER: US 08/178,302
; FILING DATE: 30-SEP-1993
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/954,148
; FILING DATE: 30-SEP-1992
; ATTORNEY/AGENT INFORMATION:
; NAME: Fitting, Thomas
; REGISTRATION NUMBER: 34,163
; REFERENCE/DOCKET NUMBER: SCR1452P
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 619-554-2937
; TELEFAX: 619-554-6312
; INFORMATION FOR SEQ ID NO: 170:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 13254 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: circular
; MOLECULE TYPE: DNA (genomic)
US-08-899-575-170

Query Match 21.4%; Score 454.8; DB 1; Length 13254;
Best Local Similarity 99.6%; Pred. No. 4e-115;
Matches 456; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1411 GTGTAGCCGTAGTGGCCACCACCTTCAGAACTCTGTAGCAGCCGCTACATACCTCGCT 1470
Db 6806 GTGTAGCCGTAGTGGCCACCACCTTCAGAACTCTGTAGCAGCCGCTACATACCTCGCT 6865
QY 1471 CTGCTAATCCTGTACCACTGCTGCTGCGGATAGCTGCTGTACCGGGTTG 1530
Db 6866 CTGCTAATCCTGTACCACTGCTGCTGCGGATAGCTGCTGTACCGGGTTG 6925
QY 1531 GACTCAAGACGATAGTTACCGGATAGCGGCGAGCGGTGCGGTGAACGGGGGTTGTCGTC 1590
Db 6926 GACTCAAGACGATAGTTACCGGATAGCGGCGAGCGGTGCGGTGAACGGGGGTTGTCGTC 6985
QY 1591 ACACAGCCAGCTTGGAGGAGCGACCTACACCGGAGCTAGATACCTACAGCGTGAGCAT 1650
Db 6986 ACACAGCCAGCTTGGAGGAGCGACCTACACCGGAGCTAGATACCTACAGCGTGAGCAT 7045
QY 1651 TGAGAAAGCCGACCTTCCCGAAGGGAGAAAGCGGAGCGGTATCCGGTAAGCGGCAGG 1710
Db 7046 TGAGAAAGCCGACCTTCCCGAAGGGAGAAAGCGGAGCGGTATCCGGTAAGCGGCAGG 7105
QY 1711 GTCGGAACAGGAGCGGACGAGGAGCTTCCAGGGGGAAACGCTGTGATCTTTATAGT 1770
Db 7106 GTCGGAACAGGAGCGGACGAGGAGCTTCCAGGGGGAAACGCTGTGATCTTTATAGT 7165
QY 1771 CTTGTGCGGTTTTCGCACTCTGACTTGAGCGCTCGATTTTCTGATGCTCGTCAGGGGG 1830
Db 7166 CTTGTGCGGTTTTCGCACTCTGACTTGAGCGCTCGATTTTCTGATGCTCGTCAGGGGG 7225
QY 1831 CGGAGCCTATGGAAGAAACCCAGCAACCGCGCGGGGG 1868
Db 7226 CGGAGCCTATGGAAGAAACCCAGCAACCGCGCGGGGG 7263

RESULT 13
US-08-899-575-156/c
; Sequence 156, Application US/08899575
; Patent No. 5804440
; GENERAL INFORMATION:
; APPLICANT: Burton, Dennis R
; APPLICANT: Barbas, Carlos F
; APPLICANT: Lerner, Richard A
; TITLE OF INVENTION: HUMAN NEUTRALIZING MONOCLONAL ANTIBODIES
; TITLE OF INVENTION: TO HUMAN IMMUNODEFICIENCY VIRUS
; NUMBER OF INVENTIONS: 170
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: The Scripps Research Institute, Office of

; ADDRESSEE: Patent Counsel
; STREET: 10666 No. 5804440th Torrey Pines Road, Suite 220,
; STREET: Mail Drop TPC8
; CITY: La Jolla
; STATE: CA
; COUNTRY: USA
; ZIP: 92037
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC Compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/899,575
; FILING DATE: 24-JUL-1997
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/276,852
; FILING DATE: 18-JUL-1994
; APPLICATION NUMBER: US 08/178,302
; FILING DATE: 30-SEP-1993
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/954,148
; FILING DATE: 30-SEP-1992
; ATTORNEY/AGENT INFORMATION:
; NAME: Fitting, Thomas
; REGISTRATION NUMBER: 34,163
; REFERENCE/DOCKET NUMBER: SCR1452P
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 619-554-2937
; TELEFAX: 619-554-6312
; INFORMATION FOR SEQ ID NO: 156:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 13254 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: circular
; MOLECULE TYPE: DNA (genomic)
US-08-899-575-156

Query Match 21.4%; Score 454.8; DB 1; Length 13254;
Best Local Similarity 99.6%; Pred. No. 4e-115;
Matches 456; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1411 GTGTAGCCGTAGTGGCCACCACCTTCAGAACTCTGTAGCAGCCGCTACATACCTCGCT 1470
Db 6449 GTGTAGCCGTAGTGGCCACCACCTTCAGAACTCTGTAGCAGCCGCTACATACCTCGCT 6390
QY 1471 CTGCTAATCCTGTACCACTGCTGCTGCGGATAGCTGCTGTACCGGGTTG 1530
Db 6389 CTGCTAATCCTGTACCACTGCTGCTGCGGATAGCTGCTGTACCGGGTTG 6330
QY 1531 GACTCAAGACGATAGTTACCGGATAGCGGCGAGCGGTGCGGTGAACGGGGGTTGTCGTC 1590
Db 6329 GACTCAAGACGATAGTTACCGGATAGCGGCGAGCGGTGCGGTGAACGGGGGTTGTCGTC 6270
QY 1591 ACACAGCCAGCTTGGAGGAGCGACCTACACCGGAGCTAGATACCTACAGCGTGAGCAT 1650
Db 6269 ACACAGCCAGCTTGGAGGAGCGACCTACACCGGAGCTAGATACCTACAGCGTGAGCAT 6210
QY 1651 TGAGAAAGCCGACCTTCCCGAAGGGAGAAAGCGGAGCGGTATCCCGTAAGCGGCAGG 1710
Db 6209 TGAGAAAGCCGACCTTCCCGAAGGGAGAAAGCGGAGCGGTATCCCGTAAGCGGCAGG 6150
QY 1711 GTCGGAACAGGAGCGGACGAGGAGCTTCCAGGGGGAAACGCTGTGATCTTTATAGT 1770
Db 6149 GTCGGAACAGGAGCGGACGAGGAGCTTCCAGGGGGAAACGCTGTGATCTTTATAGT 6090
QY 1771 CTTGTGCGGTTTTCGCACTCTGACTTGAGCGCTCGATTTTCTGATGCTCGTCAGGGGG 1830
Db 6089 CTTGTGCGGTTTTCGCACTCTGACTTGAGCGCTCGATTTTCTGATGCTCGTCAGGGGG 6030

QY 1831 CGGAGCCTATGAAACACGCGAGCAACGCGCGCGGGG 1868
Db 6029 CGGAGCCTATGAAACACGCGAGCAACGCGCGGGG 5992

RESULT 14

US-08-899-575-170
; Sequence 170, Application US/08899575
; Patent No. 5804440
; GENERAL INFORMATION:
; APPLICANT: Burton, Dennis R
; APPLICANT: Barbas, Carlos F
; APPLICANT: Lerner, Richard A
; TITLE OF INVENTION: HUMAN NEUTRALIZING MONOCLONAL ANTIBODIES
; TITLE OF INVENTION: TO HUMAN IMMUNODEFICIENCY VIRUS
; NUMBER OF SEQUENCES: 170
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: The Scripps Research Institute, Office of
; ADDRESSEE: Patent Counsel
; STREET: 10666 No. 5804440th Torrey Pines Road, Suite 220,
; CITY: La Jolla
; STATE: CA
; COUNTRY: USA
; ZIP: 92037

COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/899,575
; FILING DATE: 24-JUL-1997
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/276,852
; FILING DATE: 18-JUL-1994
; APPLICATION NUMBER: US 08/178,302
; FILING DATE: 30-SEP-1993
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/954,148
; FILING DATE: 30-SEP-1992
; ATTORNEY/AGENT INFORMATION:
; NAME: Fitting, Thomas
; REGISTRATION NUMBER: 34,163
; REFERENCE/DOCKET NUMBER: SCR1452P
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 619-554-2937
; TELEFAX: 619-554-6312
; INFORMATION FOR SEQ ID NO: 170:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 13254 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: circular
; MOLECULE TYPE: DNA (genomic)
US-08-899-575-170

Query Match 21.4%; Score 454.8; DB 1; Length 13254;
Best Local Similarity 99.6%; Pred. No. 4e-115;
Matches 456; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1411 GTGTAGCCGTAGTAGGCCACCACTTCAAGAACTCTGTAGCAGCGGCTACATACCTCGCT 1470
Db 6806 GTGTAGCCGTAGTAGGCCACCACTTCAAGAACTCTGTAGCAGCGGCTACATACCTCGCT 6865
QY 1471 CTGCTAATCTGTTACAGTGGCTGCTGCCAGTGGCGGATAGTCTTACCGGGTTG 1530
Db 6866 CTGCTAATCTGTTACAGTGGCTGCTGCCAGTGGCGGATAGTCTTACCGGGTTG 6925
QY 1531 GACTCAAGACGATAGTTACCGGATAAGGCGCAGCGGTGCGGGCTGAACGGGGGTTGCGTGC 1590

Db 6926 GACTCAAGACGATAGTTACCGGATAAGGCGCAGCGGTGCGGGCTGAACGGGGGTTGCGTGC 6985
QY 1591 ACACAGCCCGAGCTTGGAGCGAACGACCTACACCGAACTGAGATACCTACAGCGTGAGCAT 1650
Db 6986 ACACAGCCCGAGCTTGGAGCGAACGACCTACACCGAACTGAGATACCTACAGCGTGAGCAT 7045
QY 1651 TGAGAAAGCGCCACGCTTCCGAGGGGAGAAAGCGGACAGGTATCCGGTAAGCGGCGAG 1710
Db 7046 TGAGAAAGCGCCACGCTTCCGAGGGGAGAAAGCGGACAGGTATCCGGTAAGCGGCGAG 7105
QY 1711 GTGCGAAGCAGAGAGCGCACGAGGAGCTTCCAGGGGGAACGCGCTGATCTTTATAGT 1770
Db 7106 GTGCGAAGCAGAGAGCGCACGAGGAGCTTCCAGGGGGAACGCGCTGATCTTTATAGT 7165
QY 1771 CCTGTCGGGTTTCCGCCACCTCTGACTTGAGCTGATTTTGTGATGCTCGTCAGGGGGG 1830
Db 7166 CCTGTCGGGTTTCCGCCACCTCTGACTTGAGCTGATTTTGTGATGCTCGTCAGGGGGG 7225
QY 1831 CGGAGCCTATGAAACACGCGAGCAACGCGCGGGG 1868
Db 7226 CGGAGCCTATGAAACACGCGAGCAACGCGCGGGG 7263

RESULT 15

PCT-US95-08743-156/c
; Sequence 156, Application PC/TUS9508743
; GENERAL INFORMATION:
; APPLICANT:
; TITLE OF INVENTION: HUMAN NEUTRALIZING MONOCLONAL ANTIBODIES
; TITLE OF INVENTION: TO HUMAN IMMUNODEFICIENCY VIRUS
; NUMBER OF SEQUENCES: 170
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.25 (EPO)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: PCT/US95/08743
; FILING DATE: 11-JUL-1995
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/276,852
; FILING DATE: 18-JUL-1994
; INFORMATION FOR SEQ ID NO: 156:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 13254 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: circular
; MOLECULE TYPE: DNA (genomic)
PCT-US95-08743-156

Query Match 21.4%; Score 454.8; DB 5; Length 13254;
Best Local Similarity 99.6%; Pred. No. 4e-115;
Matches 456; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1411 GTGTAGCCGTAGTAGGCCACCACTTCAAGAACTCTGTAGCAGCGGCTACATACCTCGCT 1470
Db 6449 GTGTAGCCGTAGTAGGCCACCACTTCAAGAACTCTGTAGCAGCGGCTACATACCTCGCT 6390
QY 1471 CTGCTAATCTGTTACAGTGGCTGCTGCCAGTGGCGGATAGTCTTACCGGGTTG 1530
Db 6389 CTGCTAATCTGTTACAGTGGCTGCTGCCAGTGGCGGATAGTCTTACCGGGTTG 6330
QY 1531 GACTCAAGACGATAGTTACCGGATAAGGCGCAGCGGTGCGGGCTGAACGGGGGTTGCGTGC 1590
Db 6329 GACTCAAGACGATAGTTACCGGATAAGGCGCAGCGGTGCGGGCTGAACGGGGGTTGCGTGC 6270
QY 1591 ACACAGCCCGAGCTTGGAGCGAACGACCTACACCGAACTGAGATACCTACAGCGTGAGCAT 1650
Db 6269 ACACAGCCCGAGCTTGGAGCGAACGACCTACACCGAACTGAGATACCTACAGCGTGAGCAT 6210
QY 1651 TGAGAAAGCGCCACGCTTCCGAGGGGAGAAAGCGGACAGGTATCCGGTAAGCGGCGAG 1710

```
|||||
Db 6209 TGAGAAAGCGCCACGCTTCCGAAAGGAGAGAAAGGGGACAGGTATCCGGTAAAGCGGCAGG 6150
QY 1711 GTCCGAACAGGAGAGCGCACGAGGAGCTTCCAGGGGAAACGCCCTGCTATTATAGT 1770
Db 6149 GTCCGAACAGGAGAGCGCACGAGGAGCTTCCAGGGGAAACGCCCTGCTATTATAGT 6090
QY 1771 CCTGTGGGTTTCGGCACCTCTGACTTGAGCGTTCGATTTTGTGATGCTCGTCAGGGGG 1830
Db 6089 CCTGTGGGTTTCGGCACCTCTGACTTGAGCGTTCGATTTTGTGATGCTCGTCAGGGGG 6030
QY 1831 CGGAGCCTATGGAACAAACGCCAGCAACGCGCGCGGGG 1868
Db 6029 CGGAGCCTATGGAACAAACGCCAGCAACGCGCGCGGAGG 5992
```

Search completed: January 17, 2002, 11:54:39
Job time: 17406 sec

THIS PAGE BLANK (USPTO)

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: January 17, 2002, 08:54:11 ; Search time 6501.33 Seconds
(without alignments)
3512.324 Million cell updates/sec

Title: US-09-242-202a-22
Perfect score: 2125
Sequence: 1 GCCACCATGGCCCTGACCT.....AGCCTCTCCACAGGTACCT 2125

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 11351937 seqs, 537289281 residues

Total number of hits satisfying chosen parameters: 22703874

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : EST:*
1: em_estfun:*
2: em_esthum:*
3: em_estom:*
4: em_estom:*
5: em_estpl:*
6: em_estba:*
7: em_estro:*
8: em_estov:*
9: em_hic:*
10: gb_estl:*
11: gb_est2:*
12: gb_hic:*
13: gb_gss:*
14: em_gss_fun:*
15: em_gss_hum:*
16: em_gss_inv:*
17: em_gss_pln:*
18: em_gss_pro:*
19: em_gss_rod:*
20: em_gss_vrt:*
21: em_gss_other:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	507.2	23.9	685	10 AU123871	AU123871
2	481.8	22.7	621	11 BG283493	BG283493 602407782
3	479.2	22.6	791	11 B1154872	B1154872 602902857
4	455.4	21.4	649	10 AA496412	AA496412 zv37a02.r
5	450.2	21.2	615	10 AV735664	AV735664 AV735664
6	449.8	21.2	527	10 AL043585	AL043585 DKF2p434G
7	449.8	21.2	571	10 AL044178	AL044178 DKF2p434P
8	449.8	21.2	579	10 AL043613	AL043613 DKF2p434H
9	449.8	21.2	616	10 AV735756	AV735756 AV735756
10	449.8	21.2	617	10 AU281661	AU281661 4A3A-P8G1
11	449.8	21.2	628	11 BF381364	BF381364 ASIR0004
12	449.8	21.2	954	10 AL044364	AL044364 DKF2p434C

C 13	448.8	21.1	629	10	AL593919	AL593919
C 14	448.8	21.1	1004	10	AJ281480	AJ281480 4A3A-P4G8
C 15	448.4	21.1	609	10	AW701942	AW701942 uq93b02.Y
C 16	448.2	21.1	703	10	AJ281437	AJ281437 4A3A-P4C3
C 17	448.2	21.1	800	10	AJ281449	AJ281449 4A3A-P4D5
C 18	447.8	21.1	568	10	AJ281376	AJ281376 4A3A-P2G2
C 19	447.8	21.1	579	10	AJ281320	AJ281320 4A3A-P4H1
C 20	447.8	21.1	700	10	AJ281616	AJ281616 4A3A-P8A1
C 21	446.6	21.0	1067	10	AU081137	AU081137 AU081137
C 22	445	20.9	808	10	AU176264	AU176264 AU176264
C 23	443.8	20.9	548	10	AJ281654	AJ281654 4A3A-P8F1
C 24	442.8	20.8	1070	10	AJ281552	AJ281552 4A3A-P6F1
C 25	441.8	20.8	526	10	AL043840	AL043840 DKF2p434A
C 26	441.2	20.8	498	10	AL039576	AL039576 DKF2p434D
C 27	438.8	20.6	615	10	AL044413	AL044413 DKF2p434E
C 28	438.4	20.6	741	11	BF299419	BF299419 2A4-6-11
C 29	437.8	20.6	734	10	AL039459	AL039459 DKF2p434O
C 30	437.2	20.6	1089	10	AU081124	AU081124 AU081124
C 31	435.8	20.5	480	13	CT11G3	CT11G3
C 32	435.8	20.5	636	10	AJ281699	AJ281699 4A3A-P9E3
C 33	435.2	20.5	756	10	BE749097	BE749097 601123138
C 34	434.8	20.5	718	13	AG010489	AG010489 Homo sapi
C 35	433.6	20.4	959	10	BE749147	BE749147 601123355
C 36	431	20.3	759	10	BE749118	BE749118 601123394
C 37	429.6	20.2	840	10	BF863156	BF863156 601123444
C 38	429.2	20.2	757	11	BF863156	BF863156 963041F03
C 39	427.2	20.1	786	11	B1155788	B1155788 602904360
C 40	425.8	20.0	462	11	BF834727	BF834727 CM3-HT096
C 41	425.8	20.0	630	10	AL042334	AL042334 DKF2p434O
C 42	425.2	20.0	784	13	AQ876119	AQ876119 V133E1 mt
C 43	423.8	19.9	590	10	AV613078	AV613078 AV613078
C 44	422.6	19.9	795	13	AQ876011	AQ876011 V132B5 mt
C 45	418.6	19.7	690	13	AG009464	AG009464 Homo sapi

ALIGNMENTS

RESULT 1
AU123871
LOCUS AU123871 NT2RM2 Homo sapiens cDNA clone NT2RM2001211 5', mRNA
DEFINITION AU123871 sequence.
ACCESSION AU123871
VERSION AU123871.1 GI:10948587
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
AUTHORS Ota,T., Wakamatsu,A., Ozawa,M., Ishii,S., Saito,K., Yamamoto,J., Nakamura,Y., Nishikawa,T., Nagai,T., Suzuki,Y., Sugano,S. and Isogai,T.
TITLE HRI human cDNA project (Ota,T., Wakamatsu,A., Ozawa,M., Ishii,S., Saito,K., Yamamoto,J., Nakamura,Y., Nishikawa,T., Nagai,T., Suzuki,Y., Sugano,S., Isogai,T.)
JOURNAL Unpublished (2000)
COMMENT Contact: Takao Isogai
Genomics Laboratory
Helix Research Institute
1532-3 Yana, Kisarazu, Chiba 292-0812, Japan
Tel: 81-438-52-3951
Fax: 81-438-52-3952
Email: genomics@hri.co.jp
HRI human cDNA project; 5'- & 3'-end one pass sequencing: Helix Research Institute; cDNA library construction: Department of Virology, Institute of Medical Science, University of Tokyo, and Helix Research Institute.
FEATURES
source
Location/Qualifiers
1..685
/organism="Homo sapiens"
/db_xref="taxon:9606"


```

VERSION      B1154872.1  GI:14614873
KEYWORDS     EST.
SOURCE       house mouse.
ORGANISM     Mus musculus
REFERENCE    Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
AUTHORS      Mammalia; Euthera; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
TITLE        1 (bases 1 to 791)
JOURNAL      NIH-MGC http://mgc.nci.nih.gov/.
COMMENT      National Institutes of Health, Mammalian Gene Collection (MGC)
              Unpublished (1999)
              Contact: Robert Strausberg, Ph.D.
              Email: cgapbs-remail.nih.gov
              Tissue Procurement: Lohar Hennighausen Ph.D., Chu-Xia Deng Ph.D.
              cDNA Library Preparation: Life Technologies, Inc.
              cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
              DNA Sequencing by: Incyte Genomics, Inc.
              Clone distribution: MGC clone distribution information can be
              found through the I.M.A.G.E. Consortium/LLNL at:
              http://image.llnl.gov
              Plate: LLAM11090 row: b column: 15
              High quality sequence stop: 764.

FEATURES     Location/Qualifiers
             source
             1..791
             /organism="Mus musculus"
             /db_xref="taxon:10090"
             /clone="IMAGE:5032238"
             /clone_lib="NIH_CGAP_Mam3"
             /tissue_type="tumor, gross tissue"
             /lab_host="DH10B"
             /note="Organ: mammary; Vector: pCMV-SPORT6; Site_1: NotI;
             Site_2: SalI; Cloned unidirectionally. Primer: Oligo dt.
             Average insert 2 kb. Library constructed by Life
             Technologies, catalog #12017-018. Investigators providing
             samples: Lohar Hennighausen/Chu-Xia Deng, NIH Reference
             for transgenic model: Xu et al., Nature Genetics 22, 37-43
             (1999). Note: this is a NCI_CGAP Library."
BASE COUNT   159 a 264 c 215 g 153 t
ORIGIN

Query Match      22.6%; Score 479.2; DB 11; Length 791;
Best Local Similarity 81.8%; Pred. No. 3.7e-109;
Matches 589; Conservative 0; Mismatches 128; Indels 3; Gaps 3;

QY 198 GGGGTGCTGTTGGGATCCTCATCAAGCGGCGGAGAGAGATGTCACAGACCC 257
DB 11111111111111111111111111111111111111111111111111111
QY 68 GGAGCTGTGATGCTGAAGAGTACTGGTACCCAGCAGGATTTCTTCCCCAGACCC 127
DB 11111111111111111111111111111111111111111111111111111
QY 258 TGCCCCGGCGCTGGGGCATGGTCCACACAGGACCGCAGCTCATCTACCAGGAGTGG 317
DB 11111111111111111111111111111111111111111111111111111
QY 128 TGCCCTAGTACTGGGAGCACAGGCCACCGCAGACACCGCAGCTCGTCGGCCAGGAGTGG 187
QY 318 CGGTGGGACCTGACATGAGGCTGGAGCCCTCTGAAGAGAGGCCCCCAGGTCTCCACT 377
DB 11111111111111111111111111111111111111111111111111111
QY 188 CGGTGGTGAAGTGAACATGAGCTGGGCTGGAGCCCTCGGAAGAAGAGCCCGCAGATCTCCACT 247
DB 11111111111111111111111111111111111111111111111111111
QY 378 GGCACCCCTCCGAAGGGCTGGCTCGGATGTTTGTATGCTGACCTGGGAATGGGGGAGC 437
DB 11111111111111111111111111111111111111111111111111111
QY 248 GCTCCCTCCGAAGGGCTGGCTCGGATGTTTGTATGCTGACCTGGGAGTGGGGGTAAAC 307
DB 11111111111111111111111111111111111111111111111111111
QY 438 CAAGGGGTGCAAAAGCCCTCCACACATGACCCACCGCCCTCTACAGCGGTACAGTGAGGA 497
DB 11111111111111111111111111111111111111111111111111111
QY 308 CAAAGGACTGCAGAGCCCTCTCCACATGACCTCAGACCTCTACAGCGGTACAGTGAGGA 367
DB 11111111111111111111111111111111111111111111111111111
QY 498 CCCCACAGTACCCCTCTGAGCTGAGGTGACGTTCGCCCCCTTCGACCTGACCTGACGCC 557
DB 11111111111111111111111111111111111111111111111111111
QY 368 TCCACATTAACCTCTGCCCGCCGAGACTGATGGCTACGTTCCTCCCTGGCCCTGACGCC 427
DB 11111111111111111111111111111111111111111111111111111
QY 558 CCAGCTCAATATGTGAACCCAGAGTGTTCGGCCCCCAGCCCCCTTCGCCCCGAGAGGG 617
DB 11111111111111111111111111111111111111111111111111111
QY 428 CCAGCCCCGAGTATGTGAACCCAGAGGTTGGGCTCAGTCTCCCTTTGACCCCCAGAGGG 487
DB 11111111111111111111111111111111111111111111111111111
QY 618 CCCTCTGCTGTCGCCGACCTGCTGGTGCCACTCTGGAAAAGGCCCAAGACTCTCTCCCC 677

```

```

Db 488 TCCTCCGCTCCCATCCGACCTGGTGGTACTCTAGAAAGACCAAGACTCTCTCTCC 547
QY 678 AGGGAAGAAATGGGTCGTCAAAGACGTTTTCCTTTGGGGGTGCGGTGGAGAACCCCG- 736
Db 548 TGGGAAAATGGGTTGTCAAAGAGC-TTTTTCCTTTGGGGGTGCTGTGGAGAACCCCTGA 606
QY 737 AGACTTGACACCCAGGAGGAGCTGCCCTCAG-CCCCACACCTCTCTCTGCTTCAGCC 795
Db 607 ATACTTAGCACCAGCAGGACGCTCTCTCAGACCCCAACCTCTCTGCTTCAGCC 666
QY 796 CAGCCTTCGACAACTCTATTACTGGACAGCAGGACCCACAGCGGGGGTTCACACCA 855
Db 667 CAGCCTTTGACAACTCTATTACTGGACAGCAGGACCTCATCGAGCAGGCTCTCCACCAA 726
QY 856 GCACCTTCAAAGGACACCTACGGCAGAGAACCCAGACTACCTGGTCTGGAGCTGCCAG 915
Db 727 GTACCTTTGAAGGGAGCCCCCAATGACAGAGAACCTGAGTACCTAGGCTGGATGTGCCAG 786

RESULT      4
LOCUS       AA496412      649 bp      mRNA      EST      11-AUG-1997
DEFINITION zv37a02.r1 Soares ovary tumor NbHOT Homo sapiens' cDNA clone
IMAGE:755786 5' similar to gb:M11730 ERBB-2 RECEPTOR
PROTEIN-TYROSINE KINASE PRECURSOR (HUMAN);, mRNA sequence.
ACCESSION   AA496412
VERSION     AA496412.1  GI:2229733
KEYWORDS    EST.
SOURCE      human.
ORGANISM    Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Euthera; Primates; Catarrhini; Homnidae; Homo.
REFERENCE   1 (bases 1 to 649)
AUTHORS     Hillier,L., Allen,M., Bowles,L., Dubuque,T., Geisel,G., Jost,S.,
            Kucaba,T., Lacy,M., Le.N., Lennon,G., Marra,M., Martin,J., Moore,B.,
            Schellenberg,K., Steptoe,M., Tan,F., Theising,B., White,Y., Wylie
            ,T., Waterston,R. and Willson,R.
            WashU-Merck EST Project 1997
            Unpublished (1997)
            Contact: Willson RK
            Washington University School of Medicine
            4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
            Tel: 314 286 1800
            Fax: 314 286 1810
            Email: est@watson.wustl.edu
            This clone is available royalty-free through LLNL ; contact the
            IMAGE Consortium (info@image.llnl.gov) for further information.
            Seq primer: -28ml3 rev2 ET from Amersham
            High quality sequence stop: 257.
FEATURES     Location/Qualifiers
             source
             1..649
             /organism="Homo sapiens"
             /db_xref="taxon:9606"
             /clone="IMAGE:755786"
             /clone_lib="Soares ovary tumor NbHOT"
             /sex="female"
             /tissue_type="ovarian tumor"
             /lab_host="DH10B (ampicillin resistant)"
             /note="Organ: ovary; Vector: pT7T3D (Pharmacia) with a
             modified polylinker; Site_1: Not I; Site_2: Eco RI; 1st
             strand cDNA was primed with a Not I - oligo(dT) primer [5'
             TGTATACCAATCTGAAGTGGAGCGCGCGTGTGTGTGTGTGTGTGTGTGT 3'],
             double-stranded cDNA was size selected, ligated to Eco RI
             adapters (Pharmacia), digested with Not I and cloned into
             the Not I and Eco RI sites of a modified pT7T3 vector
             (Pharmacia). Library constructed by Bento Soares and
             M.Fatima Bonaldo."
BASE COUNT   145 a 202 c 185 g 116 t 1 others
ORIGIN

Query Match      21.4%; Score 455.4; DB 10; Length 649;

```


Am Klopferspitz 18a D-82152 Martinsried, Germany
This is the 3' sequence of the clone insert
Cloned from S. Wiemann, Molecular Genome Analysis, German Cancer
Research Center (DKFZ); Email: s.wiemann@dkfz-heidelberg.de;
sequenced by GBF (National Research Centre for Biotechnology Ltd.,
Braunschweig/Germany) within the cDNA sequencing consortium of the
German Genome Project.
No r1 sequence available.
This clone (DKFZp434G0127) is available at the RZPD in Berlin.
Please contact the RZPD: Ressourcenzentrum, Heubnerweg 6, 14059
Berlin-Charlottenburg, GERMANY; Email: clone@rzpd.de.

FEATURES

source
1. .527
Location/Qualifiers
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="DKFZp434G0127"
/clone_lib="434 (synonym: htes3)"
/tissue_type="testis"
/dev_stage="adult"
/lab_host="DH10B"
/note="Vector: pSport1; Site_1: NotI; Site_2: SalI"

BASE COUNT 113 a 158 c 137 g 119 t
ORIGIN
Query Match 21.2%; Score 449.8; DB 10; Length 527;
Best Local Similarity 99.6%; Pred. No. 7.5e-102;
Matches 451; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1411 GTGTAGCCGTAGTTAGGCGACACCTTCAGAACTCTGTAGCAGCGCTACATACCTCGCT 1470
Db 474 GTGTAGCCGTAGTTAGGCGACACCTTCAGAACTCTGTAGCAGCGCTACATACCTCGCT 415
QY 1471 CTGCTAATCTGTTACACGTGCTGCTCCAGTGGCGATAAGTCTTACCGGGTTG 1530
Db 414 CTGCTAATCTGTTACACGTGCTGCTCCAGTGGCGATAAGTCTTACCGGGTTG 355
QY 1531 GACTCAAGCAGTAGTTACCGGTAAGCGCGAGCGCTGCGGTGAACGGGGGTTCTGTC 1590
Db 354 GACTCAAGCAGTAGTTACCGGTAAGCGCGAGCGCTGCGGTGAACGGGGGTTCTGTC 295
QY 1591 ACACAGCCAGCTTGAGCGAGACGACCTACACGGAAGTACGATACCTACAGCTGAGCAT 1650
Db 294 ACACAGCCAGCTTGAGCGAGACGACCTACACGGAAGTACGATACCTACAGCTGAGCTA 235
QY 1651 TGAGAAAGCGCCAGCTTCCGGAAGGAGAAAGCGGAGAGTATCCGGTAAAGCGCAGG 1710
Db 234 TGAGAAAGCGCCAGCTTCCGGAAGGAGAAAGCGGAGAGTATCCGGTAAAGCGCAGG 175
QY 1711 GTCGGAACAGGAGCGCAGGAGGAGCTTCCAGGGGGAACGCCCTGTTATATAGT 1770
Db 174 GTCGGAACAGGAGCGCAGGAGGAGCTTCCAGGGGGAACGCCCTGTTATATAGT 115
QY 1771 CCTGTGGGTTTCGCCACCTCTGACTGTAGCGTCGATTTTGTGATGCTGTCAGGGGG 1830
Db 114 CCTGTGGGTTTCGCCACCTCTGACTGTAGCGTCGATTTTGTGATGCTGTCAGGGGG 55
QY 1831 CGGAGCCTATGGAAGAAACGCCAGCAGCGGCC 1863
Db 54 CGGAGCCTATGGAAGAAACGCCAGCAGCGGCC 22

RESULT 7
AL041178/c
LOCUS AL041178 571 bp mRNA EST 29-FEB-2000
DEFINITION DKFZp434P0828_s1 434 (synonym: htes3) Homo sapiens cDNA clone
ACCESSION AL041178
VERSION AL041178.3 GI:5935954
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

REFERENCE
AUTHORS
TITLE
JOURNAL
COMMENT

Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
1 (bases 1 to 571)
Bloecker,H., Boecker,M., Brandt,P., Mewes,H.W., Gassenhuber,J. and
Wiemann,S.
EST (Bloecker, et al.)
Unpublished (1999)
On Jul 9, 1999 this sequence version replaced gi:5866789.
Contact: Bloecker H
MIPS
Am Klopferspitz 18a D-82152 Martinsried, Germany
This is the 3' sequence of the clone insert
Cloned from S. Wiemann, Molecular Genome Analysis, German Cancer
Research Center (DKFZ); Email: s.wiemann@dkfz-heidelberg.de;
sequenced by GBF (National Research Centre for Biotechnology Ltd.,
Braunschweig/Germany) within the cDNA sequencing consortium of the
German Genome Project.
No r1 sequence available.
This clone (DKFZp434P0828) is available at the RZPD in Berlin.
Please contact the RZPD: Ressourcenzentrum, Heubnerweg 6, 14059
Berlin-Charlottenburg, GERMANY; Email: clone@rzpd.de.

FEATURES

Location/Qualifiers
1. .571
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="DKFZp434P0828"
/clone_lib="434 (synonym: htes3)"
/tissue_type="testis"
/dev_stage="adult"
/lab_host="DH10B"
/note="Vector: pSport1; Site_1: NotI; Site_2: SalI"
BASE COUNT 122 a 168 c 149 g 132 t
ORIGIN

Query Match 21.2%; Score 449.8; DB 10; Length 571;
Best Local Similarity 99.6%; Pred. No. 7.7e-102;
Matches 451; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1411 GTGTAGCCGTAGTTAGGCGACACCTTCAGAACTCTGTAGCAGCGCTACATACCTCGCT 1470
Db 463 GTGTAGCCGTAGTTAGGCGACACCTTCAGAACTCTGTAGCAGCGCTACATACCTCGCT 404
QY 1471 CTGCTAATCTGTTACACGTGCTGCTCCAGTGGCGATAAGTCTTACCGGGTTG 1530
Db 403 CTGCTAATCTGTTACACGTGCTGCTCCAGTGGCGATAAGTCTTACCGGGTTG 344
QY 1531 GACTCAAGCAGTAGTTACCGGTAAGCGCGAGCGCTGCGGTGAACGGGGGTTCTGTC 1590
Db 343 GACTCAAGCAGTAGTTACCGGTAAGCGCGAGCGCTGCGGTGAACGGGGGTTCTGTC 284
QY 1591 ACACAGCCAGCTTGAGCGAGACGACCTACACGGAAGTACGATACCTACAGCTGAGCAT 1650
Db 283 ACACAGCCAGCTTGAGCGAGACGACCTACACGGAAGTACGATACCTACAGCTGAGCTA 224
QY 1651 TGAGAAAGCGCCAGCTTCCGGAAGGAGAAAGCGGAGAGTATCCGGTAAAGCGCAGG 1710
Db 223 TGAGAAAGCGCCAGCTTCCGGAAGGAGAAAGCGGAGAGTATCCGGTAAAGCGCAGG 164
QY 1711 GTCGGAACAGGAGCGCAGGAGGAGCTTCCAGGGGGAACGCCCTGTTATATAGT 1770
Db 163 GTCGGAACAGGAGCGCAGGAGGAGCTTCCAGGGGGAACGCCCTGTTATATAGT 104
QY 1771 CTGTCTGGGTTTCGCCACCTCTGACTGTAGCGTCGATTTTGTGATGCTGTCAGGGGG 1830
Db 103 CTGTCTGGGTTTCGCCACCTCTGACTGTAGCGTCGATTTTGTGATGCTGTCAGGGGG 44
QY 1831 CGGAGCCTATGGAAGAAACGCCAGCAGCGGCC 1863
Db 43 CGGAGCCTATGGAAGAAACGCCAGCAGCGGCC 11

RESULT 8
AL043613/c

```
LOCUS      AL043613      579 bp      mRNA      EST      29-FEB-2000
DEFINITION DKFZp434H1527_s1.434 (synonym: htes3) Homo sapiens cDNA clone
ACCESSION  AL043613
VERSION    AL043613.1 GI:5423000
KEYWORDS   EST.
SOURCE     human.
ORGANISM   Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE  1 (bases 1 to 579)
AUTHORS   Bloecker,H., Boecker,M., Brandt,P., Mewes,H.W., Gassenhuber,J. and
            Wiemann,S.
TITLE     EST (Bloecker, et al.)
JOURNAL   Unpublished (1999)
COMMENT   Contact: Bloecker H
            MIPS
            Am Klopferapitz 18a D-82152 Martinsried, Germany
            This is the 3' sequence of the clone insert
            Clone from S. Wiemann, Molecular Genome Analysis, German Cancer
            Research Center (DKFZ); Email s.wiemann@dkfz-heidelberg.de;
            sequenced by GBF (National Research Centre for Biotechnology Ltd.,
            Braunschweig/Germany) within the cDNA sequencing consortium of the
            German Genome Project.
            No r1 sequence available.
            This clone (DKFZp434H1527) is available at the RZPD in Berlin.
            Please contact the RZPD: Ressourcenzentrum, Heubnerweg 6, 14059
            Berlin-Charlottenburg, GERMANY; Email: clone@rzpd.de.
FEATURES   Location/Qualifiers
            source          1..579
                        /organism="Homo sapiens"
                        /db_xref="taxon:9606"
                        /clone_lib="434 (synonym: htes3)"
                        /tissue_type="testis"
                        /dev_stage="adult"
                        /lab_host="pH10B"
                        /note="Vector: pSPort1; Site_1: NotI; Site_2: SalI"
BASE COUNT 128 a 171 c 152 g 128 t
ORIGIN
Query Match      21.2%; Score 449.8; DB 10; Length 579;
Best Local Similarity 99.6%; Pred. No. 7.7e-102;
Matches 451; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
QY 1411 GTGTAGCCCTAGTTAGGCCACCACTTCAAGAACTCTGTAGCACCGCTACATACCTCGCT 1470
DB 477 GTGTAGCCCTAGTTAGGCCACCACTTCAAGAACTCTGTAGCACCGCTACATACCTCGCT 418
QY 1471 CTGCTAATCTCTTACCACTGCTGCTGCCAGTGGCGATAAGTCTGTGCTTACCGGGTTG 1530
DB 417 CTGCTAATCTCTTACCACTGCTGCTGCCAGTGGCGATAAGTCTGTGCTTACCGGGTTG 358
QY 1531 GACTCAAGCAGTAGTTACCGGATAAGCGGACGGCTCGGGCTGAACGGGGGTTTCGTGC 1590
DB 357 GACTCAAGCAGTAGTTACCGGATAAGCGGACGGCTCGGGCTGAACGGGGGTTTCGTGC 298
QY 1591 ACACAGCCCCAGTTGAGCGAAGCACTTACACCGAATCGAGATACCTACACGCTGAGCAT 1650
DB 297 ACACAGCCCCAGTTGAGCGAAGCACTTACACCGAATCGAGATACCTACACGCTGAGCTA 238
QY 1651 TGAGAAAGCGCCACCTTCCGAGAGGAGAAAGCGGACAGGTATCCGGTAAGCGGCAGG 1710
DB 237 TGAGAAAGCGCCACCTTCCGAGAGGAGAAAGCGGACAGGTATCCGGTAAGCGGCAGG 178
QY 1711 GTCGAAGCAGGAGCGCACGAGGAGCTTCCAGGGGAAACCGCTGTATCTTTATAGT 1770
DB 177 GTCGAAGCAGGAGCGCACGAGGAGCTTCCAGGGGAAACCGCTGTATCTTTATAGT 118
QY 1771 CCTGTCGGGTTTCGCCACCTCTGACTTGAGCGTCGATTTTGTGATCGTCTGTCAGGGGG 1830
DB 117 CCTGTCGGGTTTCGCCACCTCTGACTTGAGCGTCGATTTTGTGATCGTCTGTCAGGGGG 58
```

```
QY 1831 CGGAGCCTATGGAAAAACGCCAGCAACGCGGCC 1863
DB 57 CGGAGCCTATGGAAAAACGCCAGCAACGCGGCC 25

RESULT 9
AV735756/c
LOCUS      AV735756      616 bp      mRNA      EST      17-OCT-2000
DEFINITION AV735756 CB Homo sapiens cDNA clone CBMAGC03 5', mRNA sequence.
ACCESSION  AV735756
VERSION    AV735756.1 GI:10853337
KEYWORDS   EST.
SOURCE     human.
ORGANISM   Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE  1 (bases 1 to 616)
AUTHORS   Zhang,Q., Ye,M., Wu,X., Gu,J., Huang,Q., Zhou,J., Shen,Y., Han,Z.,
            Chen,S., Mao,M. and Chen,Z.
TITLE     Homo sapiens CB library cDNA clones
JOURNAL   Unpublished (2000)
COMMENT   Contact: Zhu Chen
            Shanghai Institute of Hematology, Rui-Jin Hospital
            197 Rui-jin II Road, Shanghai 200025, P. R. China
            Tel: 86-21-64740490
            Fax: 86-21-64743206
            Email: mbs@shims.sh.cn
            This clone is available at Shanghai Hematology Institute in
            Shanghai.
            Chinese National Human Genome Center at Shanghai
            351 Guo Shoujing Road, Zhangjiang Hi-Tech Park, Pudong.
FEATURES   Location/Qualifiers
            source          1..616
                        /organism="Homo sapiens"
                        /db_xref="taxon:9606"
                        /clone_lib="CBMAGC03"
                        /tissue_type="CB"
                        /tissue_type="cord blood"
                        /cell_type="CD34+ hematopoietic stem/progenitor cell"
                        /lab_host="BM25.8"
                        /note="Vector: pBluescript; Site_1: EcoRI; The insert is
                        cloned randomly with the EcoRI digestion"
BASE COUNT 125 a 182 c 165 g 143 t 1 others
ORIGIN
Query Match      21.2%; Score 449.8; DB 10; Length 616;
Best Local Similarity 99.6%; Pred. No. 7.8e-102;
Matches 451; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
QY 1411 GTGTAGCCCTAGTTAGGCCACCACTTCAAGAACTCTGTAGCACCGCTACATACCTCGCT 1470
DB 487 GTGTAGCCCTAGTTAGGCCACCACTTCAAGAACTCTGTAGCACCGCTACATACCTCGCT 428
QY 1471 CTGCTAATCTCTTACCACTGCTGCTGCCAGTGGCGATAAGTCTGTCTTACCGGGTTG 1530
DB 427 CTGCTAATCTCTTACCACTGCTGCTGCCAGTGGCGATAAGTCTGTCTTACCGGGTTG 368
QY 1531 GACTCAAGCAGTAGTTACCGGATAAGCGGACGGCTCGGGCTGAACGGGGGTTTCGTGC 1590
DB 367 GACTCAAGCAGTAGTTACCGGATAAGCGGACGGCTCGGGCTGAACGGGGGTTTCGTGC 308
QY 1591 ACACAGCCCCAGTTGAGCGAAGCACTTACACCGAATCGAGATACCTACACGCTGAGCAT 1650
DB 307 ACACAGCCCCAGTTGAGCGAAGCACTTACACCGAATCGAGATACCTACACGCTGAGCTA 248
QY 1651 TGAGAAAGCGCCACCTTCCGAGAGGAGAAAGCGGACAGGTATCCGGTAAGCGGCAGG 1710
DB 247 TGAGAAAGCGCCACCTTCCGAGAGGAGAAAGCGGACAGGTATCCGGTAAGCGGCAGG 188
QY 1711 GTCGAAGCAGGAGCGCACGAGGAGCTTCCAGGGGAAACCGCTGTATCTTTATAGT 1770
```

```

Db 187 GTCCGAACAGGAGCGCACGAGGAGGCTTCCAGGGGAAACGCGCTGGTATCTTTATAGT 128
QY 1771 CTTGTCTGGGTTTCGCCACCTCTGACTTCTGAGCGTCGATTTTCTGTATCTCTGTCAGSGGG 1830
Db 127 CTTGTCTGGGTTTCGCCACCTCTGACTTCTGAGCGTCGATTTTGTGTGCTCTGTCAGGGGG 68
QY 1831 CGGAGCCTATGGAAGAACGCCAGCAACGCGGCC 1863
Db 67 CGGAGCCTATGGAAGAACGCCAGCAACGCGGCC 35

RESULT 10
AJ281661/c 617 bp mRNA EST 30-JUN-2000
LOCUS 4A3A-P8G10-F Anopheles gambiae immune competent 4A3A Anopheles
DEFINITION gambiae cDNA clone 4A3A-P8G10, mRNA sequence.
ACCESSION AJ281661
VERSION AJ281661.1 GI:5929540
KEYWORDS EST.
SOURCE African malaria mosquito.
ORGANISM Anopheles gambiae
Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
Pterygota; Neoptera; Endopterygota; Diptera; Nematocera; Culicoidea
; Anopheles.
REFERENCE 1 (bases 1 to 617)
AUTHORS Dimopoulos, G., Casavant, T.L., Chang, S., Scheetz, T., Roberts, C.,
Donohue, M., Schultz, J., Benes, V., Bork, P., Ansorge, W., Soares, M.B.
and Kafatos, F.C.
TITLE Anopheles gambiae pilot gene discovery project: identification of
mosquito innate immunity genes from expressed sequence tags
generated from immune-competent cell lines
Proc. Natl. Acad. Sci. U.S.A. 97 (12), 6619-6624 (2000)
20300950
Contact: Dimopoulos G
Fotis C. Kafatos laboratory
European Molecular Biology Laboratory
Meyerohofstrasse 1, 69117 Heidelberg, Germany.
JOURNAL Location/Qualifiers
MEDLINE 1. 617
COMMENT /organism="Anopheles gambiae"
/strain="4A r/r"
/db_xref="taxon:7165"
/clone="4A3A-P8G10"
/clone_lib="Anopheles gambiae immune competent 4A3A"
/cell_line="immune competent 4A3A"
/lab_host="E. coli DH10B"
/note="Vector: p7T3D-Pac (Pharmacia) with a modified
polylinker; Site_1: EcoRI; Site_2: NotI; sequenced from
forward priming site which reads from the 3' end of the
cDNA. The 4A3A is a directionally cloned and normalized
cDNA library that was constructed from the 4A3A cell line
oligo-T primed cDNA according to: Bonaldo, Lennon & Soares
(1996) : Normalization and Subtraction: Two approaches to
Facilitate Gene Discovery, Genome Research 6, 791-806."
BASE COUNT 136 a 175 c 162 g 144 t
ORIGIN

Query Match 21.2%; Score 449.8; DB 10; Length 617;
Best Local Similarity 99.6%; Pred. No. 7.8e-102;
Matches 451; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1411 GTGTAGCGGTAGTTAGGCACCACTTCAAGAACTCTGTAGACCCGCCCTACATACCTCGCT 1470
Db 488 GTGTAGCGGTAGTTAGGCACCACTTCAAGAACTCTGTAGACCCGCCCTACATACCTCGCT 429
QY 1471 CTGCTAACTCTTACCAGTGGCTGCTGCCAGTGGCGGATAAGTCTGTCTTACCGGGTTG 1530
Db 428 CTGCTAACTCTTACCAGTGGCTGCTGCCAGTGGCGGATAAGTCTGTCTTACCGGGTTG 369
QY 1531 GACTCAAGACGATAGTTACCGGATAAGCGCGCGGTGGGGCTGAACGGGGGGTTTCGTGC 1590
Db 368 GACTCAAGACGATAGTTACCGGATAAGCGCGCGGTGGGGCTGAACGGGGGGTTTCGTGC 309

```

```

QY 1591 ACACAGCCCGAGCTTGGAGCGAACGACCTACACCGAACTGAGATACCTACAGCGTGAGCAT 1650
Db 308 ACACAGCCCGAGCTTGGAGCGAACGACCTACACCGAACTGAGATACCTACAGCGTGAGCTA 249
QY 1651 TGAGAAAGCGCCACGCTTCCCGAAGGGAGAGAAAGCGGACAGGTATCCGGTAAGCGCGCAGG 1710
Db 248 TGAGAAAGCGCCACGCTTCCCGAAGGGAGAGAAAGCGGACAGGTATCCGGTAAGCGCGCAGG 189
QY 1711 GTCGGAACAGGAGCGCACGAGGAGGCTTCCAGGGGGAACGCGCTGTATCTTTATAGT 1770
Db 188 GTCGGAACAGGAGCGCACGAGGAGGCTTCCAGGGGGAACGCGCTGTATCTTTATAGT 129
QY 1771 CTTGTCTGGGTTTCGCCACCTCTGACTTCTGAGCGTCGATTTTCTGTATCTCTGTCAGSGGG 1830
Db 128 CTTGTCTGGGTTTCGCCACCTCTGACTTCTGAGCGTCGATTTTCTGTATCTCTGTCAGSGGG 69
QY 1831 CGGAGCCTATGGAAGAACGCCAGCAACGCGGCC 1863
Db 68 CGGAGCCTATGGAAGAACGCCAGCAACGCGGCC 36

RESULT 11
BF381364/c 628 bp mRNA EST 27-NOV-2000
LOCUS BF381364
DEFINITION ASTRO004 Mosquito Genes Pool Related Malaria Infection Anopheles
stephensi cDNA 5', mRNA sequence.
ACCESSION BF381364
VERSION BF381364.1 GI:11370487
KEYWORDS EST.
SOURCE Anopheles stephensi.
ORGANISM Anopheles stephensi.
Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
Pterygota; Neoptera; Endopterygota; Diptera; Nematocera; Culicoidea
; Anopheles.
REFERENCE 1 (bases 1 to 628)
AUTHORS Xu, X., Qu, F., Song, G. and Xu, J.
TITLE The differentially expressing genes pool from Anopheles stephensi
related to infection with Plasmodium yoelii enriched by suppression
subtractive hybridization
Unpublished (2001)
JOURNAL Contact: Xu Xiaochun; Qu Fengyi; Song Guan hong; Xu Jiannong
COMMENT Department of Parasitology
Second Military Medical University
800 Xiangyin Rd., Shanghai, 200433, China
Tel: 86 021 25070276
Email: xcxu@mmu.edu.cn
Seq primer: T7
High quality sequence stop: 629
POLYA=No.
FEATURES Location/Qualifiers
source 1. 628
/organism="Anopheles stephensi"
/strain="Hor"
/db_xref="taxon:30069"
/clone_lib="Mosquito Genes Pool Related Malaria Infection"
/sex="female"
/tissue_type="whole body"
/dev_stage="24 hours post-infection"
BASE COUNT 145 a 178 c 169 g 136 t
ORIGIN

Query Match 21.2%; Score 449.8; DB 11; Length 628;
Best Local Similarity 99.6%; Pred. No. 7.9e-102;
Matches 451; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1411 GTGTAGCGGTAGTTAGGCACCACTTCAAGAACTCTGTAGACCCGCCCTACATACCTCGCT 1470
Db 569 GTGTAGCGGTAGTTAGGCACCACTTCAAGAACTCTGTAGACCCGCCCTACATACCTCGCT 510
QY 1471 CTGCTAACTCTTACCAGTGGCTGCTGCCAGTGGCGGATAAGTCTGTCTTACCGGGTTG 1530

```

Db	508	CTGCTAACTCTGTTACCACTGGCTGTCGCGATAAGTCGTCTTACCGGGTTG	450
QY	1531	GACTCAAGACGATAGTTACCGGATAAGCGCAGCGTCGGCTGAACGGGGGTTCGTGC	1590
Db	449	GACTCAAGACGATAGTTACCGGATAGCGGCAGCGCTGGAACGGGGGTTTCGTGC	390
QY	1591	ACACAGCCCACTTTGGAGCGGAACGACCTACACCGAACTTGAGATACCTACAGCCTGAGCAT	1650
Db	389	ACACAGCCCACTTTGGAGCGGAACGACCTACACCGAACTTGAGATACCTACAGCCTGAGCTA	330
QY	1651	TGAGAAGCGCCACGCTTCCCAAGGGAGAAAGCGGACAGGTATCCGTTAAGCGGCAGG	1710
Db	329	TGAGAAGCGCCACGCTTCCCAAGGGAGAAAGCGGACAGGTATCCGTTAAGCGGCAGG	270
QY	1711	GTCCGAACAGGAGAGCGCACAGGAGGAGCTTCCAGGGGAAACGCTCGGTATCTTTATAGT	1770
Db	269	GTCCGAACAGGAGAGCGCACAGGAGGAGCTTCCAGGGGAAACGCTCGGTATCTTTATAGT	210
QY	1771	CTGTGCGGTTTTCGCAACCTCTGACTTTCGAGCGTCGATTTTTGTGATCTCGTCAGGGGGG	1830
Db	209	CTGTGCGGTTTTCGCAACCTCTGACTTTCGAGCGTCGATTTTTGTGATCTCGTCAGGGGGG	150
QY	1831	CGGAGCGCTATGGAAAAACCCACGAAACCGGCC	1863
Db	149	CGGAGCGCTATGGAAAAACCCACGAAACCGGCC	117

RESULT	12
AL044364/c	
LOCUS	954 bp mRNA EST 29-FEB-2000
DEFINITION	DKFZp43AC172_s1_434 (synonym: htes3) Homo sapiens cDNA clone
ACCESSION	DKFZp43AC172 3', mRNA sequence.
VERSION	AL044364
KEYWORDS	AL044364.1 GI:5432586
SOURCE	EST.
ORGANISM	human. Homo sapiens Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo. 1 (bases 1 to 954) Wiesang,W., Benes,V., Krieger,S., Mewes,H.W., Gassenhuber,J. and Wiemann,S.
REFERENCE	EST (Ansorge, Benes, et al.) Unpublished (1999)
AUTHORS	Contact: Ansorge W
TITLE	MIPS
JOURNAL	Am Klopferspitz 18a D-82152 Martinsried, Germany
COMMENT	This is the 3' sequence of the clone insert Clone from S. Wiemann, Molecular Genome Analysis, German Cancer Research Center (DKFZ); Email s.wiemann@kfz-heidelberg.de; sequenced by EMBL (European Molecular Biology Laboratories, Heidelberg/Germany) within the cDNA sequencing consortium of the German Genome Project. No r1 sequence available. This clone (DKFZp43AC172) is available at the RZPD in Berlin. Please contact the RZPD: Ressourcenzentrum, Heubnerweg 6, 14059 Berlin-Charlottenburg, GERMANY; Email: clone@rzpd.de.

```

FEATURES
  source
    1. .954
      Location/Qualifiers
        organism="Homo sapiens"
        db_xref="taxon:9606"
        clone="DXFZp434C172"
        clone_lib="434 (synonym: htes3)"
        tissue_type="testis"
        dev_stage="adult"
        lab_host="DH10B"
        note="Vector: pSpot11; Site_1: NotI; Site_2: SalI"
      214 a 273 c 253 g 214 t
ORIGIN
Query Match 21.2%; Score 449.8; DB 10; Length 954;
Best Local Similarity 99.6%; Pred. No. 8.9e-102;

```

Matches	451;	Conservative	0;	Mismatches	2;	Indels	0;	Gaps	0
QY	1411	GTGTAGCGCTAGTGTAGGCGCACCACTTCAAGAACTCTGTGTAGCAGCGCTACATACCTCGCT	1470						
Db	851	GTGTAGCGCTAGTGTAGGCGCACCACTTCAAGAACTCTGTGTAGCAGCGCTACATACCTCGCT	792						
QY	1471	CTGTAAATCCTGTTACCAAGTGCTGCTGCCAGTGGCGATAAGTCTGTCTTTACCGGGTTG	1530						
Db	791	CTGTAAATCCTGTTACCAAGTGCTGCTGCCAGTGGCGATAAGTCTGTCTTTACCGGGTTG	732						
QY	1531	GACTCAAGACGATAGTTCACCGGATAAGCGGACGCTGGCGCTGAACGGGGGTTCTGTC	1590						
Db	731	GACTCAAGACGATAGTTCACCGGATAAGCGGACGCTGGCGCTGAACGGGGGTTCTGTC	672						
QY	1591	ACACAGCCCCAGCTTGGAGCGAACGCTACACCGAACTGAGATACCTACAGCGTGAGCAT	1650						
Db	671	ACACAGCCCCAGCTTGGAGCGAACGCTACACCGAACTGAGATACCTACAGCGTGAGCAT	612						
QY	1651	TGAGAAAGCGCCACGCTTCCGGAAGGGAGAAAGCGGACAGATATCCGGTAAAGCGGAGG	1710						
Db	611	TGAGAAAGCGCCACGCTTCCGGAAGGGAGAAAGCGGACAGATATCCGGTAAAGCGGAGG	552						
QY	1711	GTGGAACAGGAGAGCGACGAGGAGCTTCCAGGGGAAACGCCCTGTTATCTTTATAGT	1770						
Db	551	GTGGAACAGGAGAGCGACGAGGAGCTTCCAGGGGAAACGCCCTGTTATCTTTATAGT	492						
QY	1771	CCTGTGCGGTTTCCGACCTCTGTACTTGAGCGTGATTTTGTGTGATCTCTCAGGGGGG	1830						
Db	491	CCTGTGCGGTTTCCGACCTCTGTACTTGAGCGTGATTTTGTGTGATCTCTCAGGGGGG	432						
QY	1831	CGGAGCCTATGAAAAAGCCAGCAACCGGCC	1863						
Db	431	CGGAGCCTATGAAAAAGCCAGCAACCGGCC	399						
RESULT	13								
LOCUS	AL593919/c								
DEFINITION	AL593919 XGC-gastrula Silurana tropicalis	629 bp	mRNA	EST	30-JUL-2001				
ACCSSION	AL593919								
VERSION	AL593919.1	GI:15005980							
KEYWORDS	EST.								
SOURCE	western clawed frog.								
ORGANISM	Silurana tropicalis								
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;								
AUTHORS	Amphibia; Batrachia; Anura; Mesobatrachia; Pipidae;								
	Xenopodinae; Silurana.								
	1 (bases 1 to 629)								
	Huckle,E., Taylor,R., McMurray,A., Ashurst,J.L., Zorn,A.M. and								
	Rogers,J.								
TITLE	Sanger Xenopus tropicalis EST project 2001								
JOURNAL	Unpublished (2001)								
COMMENT	Contact: Huckle E								
	Sanger Centre								
	Hinxton, Cambridgeshire, CB10 1SA, UK								
	Email: trop@sanger.ac.uk								
	Sanger Xenopus tropicalis EST project 2001								
	TROPICALIS_SEQUENCE_ID: TGas003o10.sp6								
	Sequencing primer: SP6								
	This sequence is from a Xenopus Gene Collection (XGC) library								
	constructed by Aaron M. Zorn.								
FEATURES	Location/Qualifiers								
source	1..629								

```

FEATURES
    source
        constructed by Aaron M. Zorn.
        Location/Qualifiers
            1..629
                /organism="Silurana tropicalis"
                /db_xref="taxon:8364"
                /clone="RGAS003o10"
                /clone_lib="XGC-gastrula"
                /dev_stage="gastrula (stages 10.5-13 mixed)"
                /lab_host="Escherichia coli DH10B"
                /note="Vector: pCS107; Site_1: EcoRI; Site_2: NotI; cDNA
                    was oligo d primed from 5ug of poly A+ RNA from stages
                    10-13 gastrulae. EcoRI-NotI cut cDNA was then ligated

```

Query Match 21.2%; Score 449.8; DB 10; Length 954;
Best Local Similarity 99.6%; Pred. No. 8.9e-102;

THIS PAGE BLANK (USPTO)

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: January 17, 2002, 12:03:53 ; Search time 599.86 Seconds
(without alignments)
17.151 Million cell updates/sec

Title: US-09-242-202a-22_copy_1_12
Perfect score: 12
Sequence: 1 GCCACCATGGCC 12

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 930621 seqs, 428662619 residues

Total number of hits satisfying chosen parameters: 1861242

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : N_Geneseq_1101.*
1: /SIDS2/gcgdata/geneseq/geneseq/NA1980.DAT.*
2: /SIDS2/gcgdata/geneseq/geneseq/NA1981.DAT.*
3: /SIDS2/gcgdata/geneseq/geneseq/NA1982.DAT.*
4: /SIDS2/gcgdata/geneseq/geneseq/NA1983.DAT.*
5: /SIDS2/gcgdata/geneseq/geneseq/NA1984.DAT.*
6: /SIDS2/gcgdata/geneseq/geneseq/NA1985.DAT.*
7: /SIDS2/gcgdata/geneseq/geneseq/NA1986.DAT.*
8: /SIDS2/gcgdata/geneseq/geneseq/NA1987.DAT.*
9: /SIDS2/gcgdata/geneseq/geneseq/NA1988.DAT.*
10: /SIDS2/gcgdata/geneseq/geneseq/NA1989.DAT.*
11: /SIDS2/gcgdata/geneseq/geneseq/NA1990.DAT.*
12: /SIDS2/gcgdata/geneseq/geneseq/NA1991.DAT.*
13: /SIDS2/gcgdata/geneseq/geneseq/NA1992.DAT.*
14: /SIDS2/gcgdata/geneseq/geneseq/NA1993.DAT.*
15: /SIDS2/gcgdata/geneseq/geneseq/NA1994.DAT.*
16: /SIDS2/gcgdata/geneseq/geneseq/NA1995.DAT.*
17: /SIDS2/gcgdata/geneseq/geneseq/NA1996.DAT.*
18: /SIDS2/gcgdata/geneseq/geneseq/NA1997.DAT.*
19: /SIDS2/gcgdata/geneseq/geneseq/NA1998.DAT.*
20: /SIDS2/gcgdata/geneseq/geneseq/NA1999.DAT.*
21: /SIDS2/gcgdata/geneseq/geneseq/NA2000.DAT.*
22: /SIDS2/gcgdata/geneseq/geneseq/NA2001.DAT.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	12	100.0	12	19	AAV21735
2	12	100.0	19	21	AAZ51587
3	12	100.0	19	21	AAZ51588
4	12	100.0	23	22	AAZ51589
5	12	100.0	23	22	AAZ51590
6	12	100.0	23	22	AAZ51591
7	12	100.0	25	14	AAQ41717
8	12	100.0	25	14	AAQ317109
9	12	100.0	27	14	AAQ52485
10	12	100.0	27	14	AAQ67276
11	12	100.0	27	16	AAQ1078
12	100.0	27	16	AAQ97620	

12	12	100.0	27	18	AAT84743	Primer BPI-23 for
13	12	100.0	27	19	AAV21728	P53 cDNA RT-PCR pr
14	12	100.0	28	20	AAZ28010	Human lectomedin-2
15	12	100.0	28	21	AAA96037	Human lectomedin-2
16	12	100.0	30	22	AAZ83005	Human MBSP2 cDNA c
17	12	100.0	31	20	AAV90571	Forward PCR primer
18	12	100.0	32	21	AAZ11968	Human truncated pI
19	12	100.0	32	22	AAZ77491	Alphalib integrin
20	12	100.0	33	17	AAT17090	Murine class II MH
21	12	100.0	33	19	AAZ62945	Murine signalling
22	12	100.0	33	20	AAZ02029	Murine signalling
23	12	100.0	33	21	AAZ60478	Murine signalling
24	12	100.0	33	21	AAA30367	Plasmid TKHH2 PCR
25	12	100.0	33	21	AAZ14288	Murine MHC class I
26	12	100.0	33	21	AAZ86757	PCR primer #6048 f
27	12	100.0	33	21	AAZ38665	Ig-betaTh fusion g
28	12	100.0	33	22	AAZ27536	Primer #1. Uniden
29	12	100.0	33	22	AAZ83456	Primer #1 used in
30	12	100.0	34	17	AAT09016	FLAP ifisense prim
31	12	100.0	35	21	AAZ63201	Adenovirus E2A cod
32	12	100.0	35	21	AAZ63004	Wildtype early reg
33	12	100.0	35	21	AAZ52587	Adenovirus 5 E2A r
34	12	100.0	35	22	AAZ80211	PCR primer used to
35	12	100.0	35	22	AAZ55483	PCR primer used to
36	12	100.0	35	22	AAZ58071	Adenovirus E2A cod
37	12	100.0	35	22	AAZ88882	Adenovirus E2A pri
38	12	100.0	36	19	AAZ17110	Glutathione S-tran
39	12	100.0	36	19	AAZ17111	Glutathione S-tran
40	12	100.0	36	22	AAZ56642	Human alpha interf
41	12	100.0	36	22	AAZ28415	Human TGC cDNA PCR
42	12	100.0	37	22	AAZ32571	Asp2 coding sequen
43	12	100.0	38	19	AAZ32758	Human interferon-a
44	12	100.0	38	22	AAZ75446	Codon-optimised HP
45	12	100.0	39	21	AAZ49559	Human delta-5-desa

ALIGNMENTS

RESULT 1
AAV21735
ID AAV21735 standard; cDNA; 12 BP.
XX
AC AAV21735;
XX
DT 17-AUG-1998 (first entry)
XX
DE Humanised vector 5' sequence acceptance site.
XX
KW Vector; vaccine; tumour; antigen; plasmid pITL; ss.
XX
OS Synthetic.
XX
PN WO9806863-A1.
XX
PD 19-FEB-1998.
XX
PF 14-AUG-1997; 97WO-US14306.
XX
PR 14-AUG-1996; 96US-0023931.
XX
(USSH) US DEPT HEALTH & HUMAN SERVICES.
XX
PI Nelson EL, Nelson PJ;
XX
DR WPI; 1998-159552/14.
XX
PT Humanised polynucleotide vectors - comprising human derived promoter
PT and sequence acceptance site, used for the production of vaccines
XX
PS Claim 13; Page 104; 125pp; English.
XX
CC This DNA sequence comprises a 5' acceptance site of novel humanised

CC polynucleotide vectors such as pITL (see AAV21724). Such vectors
 CC comprise a human-derived promoter or mammalian homologue and a
 CC sequence acceptance site that is designed to directionally accept
 CC sequence specific products from RT-PCR based cloning strategies
 CC via a unique site within an interrupted palindrome recognition
 CC sequence for a restriction endonuclease which is incorporated into
 CC the PCR primer. In this embodiment, the palindrome recognition
 CC sequence is for BglI. The 5' acceptance site reads on the positive
 CC strand, and includes a GCC sequence recognised by BglI and an ATG
 CC start codon. The 3' GCC moiety keeps the sequence in-frame and
 CC encodes Ser. The 3' acceptance site is given in AAV21736. The novel
 CC vectors are used to express target antigens, especially tumour
 CC antigens. They are non-replicating in mammalian cells but are
 CC capable of extended stable expression of target sequences, and
 CC generate immune responses to the target sequences with little or no
 CC immune response to the other vector components.

XX
 SQ Sequence 12 BP; 2 A; 6 C; 3 G; 1 T; 0 other;

Query Match 100.0%; Score 12; DB 19; Length 12;
 Best Local Similarity 100.0%; Pred. No. 7.9e+02;
 Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GCCACCATGGCC 12
 Db 1 gccaccatggcc 12

RESULT 2
 AAZ51587
 ID AAZ51587 standard; DNA; 19 BP.
 XX
 AC AAZ51587;
 XX
 DT 21-JUN-2000 (first entry)
 XX
 DE PCR primer 5'ORF1 to amplify human Deltex (hZDX) coding region.
 XX
 KW Notch signalling protein; Deltex; hZDX; human; cell differentiation;
 KW proliferation; ankyrin repeat; regulator; cytoskeletal; modulator; cancer;
 KW diagnosis; treatment; T lymphoblastic lymphoma; leukemia; lung carcinoma;
 KW adenocarcinoma; cervical neoplasia; PCR primer; ss.
 XX
 OS Homo sapiens.
 XX
 PN WO200012698-A1.
 XX
 PD 09-MAR-2000.
 XX
 PF 25-AUG-1999; 99WO-GB02802.
 XX
 PR 31-AUG-1998; 98US-0098512.
 XX
 PA (ZENE) ZENECA LTD.
 XX
 PI Khoury-Christianson AM, Kays JS;
 XX
 DR WPI; 2000-256639/22.
 XX

Novel human cytoplasmic Notch signaling protein and polynucleotide
 PT useful for identifying compounds that modulate Notch signaling protein
 PT and for diagnosis, treatment of leukemia, T lymphoblastic lymphoma -
 XX
 PS Example 2; Page 16; 60pp; English.

CC The present DNA sequence is the PCR primer, 5'ORF1 used to amplify the
 CC coding region from the start site of ORF1 through the common stop codon
 CC (479-2347), using human spleen cDNA library as the template. The 5' end
 CC PCR primer incorporates a Kozak consensus sequence prior to the ATG
 CC codon. The human cytoplasmic Notch signalling protein, Deltex (hZDX),
 CC is integral to cell differentiation and proliferation. It binds to the
 CC ankyrin repeat region of human Notch receptor and functions as a

CC regulator of Notch signalling pathway and has cytostatic activity. hZDX
 CC sequence is useful to identify compounds that modulate biological or
 CC pharmacological activity of Deltex Notch signalling protein. Specific
 CC antisense oligomers and dominant negative mutants are useful for
 CC therapeutic purposes. It is also used for diagnosis and treatment of
 CC conditions associated with aberrant Notch signalling, like cancers,
 CC including T lymphoblastic lymphoma, leukemia, adenocarcinoma, lung
 CC carcinoma and cervical neoplasia. The nucleotide sequence is useful
 CC for diagnostic assays to detect expression levels of hZDX.

XX Sequence 19 BP; 3 A; 9 C; 5 G; 2 T; 0 other;

Query Match 100.0%; Score 12; DB 21; Length 19;
 Best Local Similarity 100.0%; Pred. No. 8e+02;
 Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GCCACCATGGCC 12
 Db 1 gccaccatggcc 12

RESULT 3
 AAZ51588
 ID AAZ51588 standard; DNA; 19 BP.
 XX
 AC AAZ51588;
 XX
 DT 21-JUN-2000 (first entry)
 XX
 DE PCR primer 5'ORF2 to amplify human Deltex (hZDX) coding region.
 XX
 KW Notch signalling protein; Deltex; hZDX; human; cell differentiation;
 KW proliferation; ankyrin repeat; regulator; cytoskeletal; modulator; cancer;
 KW diagnosis; treatment; T lymphoblastic lymphoma; leukemia; lung carcinoma;
 KW adenocarcinoma; cervical neoplasia; PCR primer; ss.
 XX
 OS Homo sapiens.
 XX
 PN WO200012698-A1.
 XX
 PD 09-MAR-2000.
 XX
 PF 25-AUG-1999; 99WO-GB02802.
 XX
 PR 31-AUG-1998; 98US-0098512.
 XX
 PA (ZENE) ZENECA LTD.
 XX
 PI Khoury-Christianson AM, Kays JS;
 XX
 DR WPI; 2000-256639/22.
 XX

Novel human cytoplasmic Notch signaling protein and polynucleotide
 PT useful for identifying compounds that modulate Notch signaling protein
 PT and for diagnosis, treatment of leukemia, T lymphoblastic lymphoma -
 XX
 PS Disclosure; Page 16; 60pp; English.

CC The present DNA sequence is the PCR primer, 5'ORF2 used to amplify the
 CC coding region from the start site of ORF2 through the common stop codon
 CC (485-2347), using human spleen cDNA library as the template. The 5' end
 CC PCR primer incorporates a Kozak consensus sequence prior to the ATG
 CC codon. The human cytoplasmic Notch signalling protein, Deltex (hZDX),
 CC is integral to cell differentiation and proliferation. It binds to the
 CC ankyrin repeat region of human Notch receptor and functions as a
 CC regulator of Notch signalling pathway and has cytostatic activity. hZDX
 CC sequence is useful to identify compounds that modulate biological or
 CC pharmacological activity of Deltex Notch signalling protein. Specific
 CC antisense oligomers and dominant negative mutants are useful for
 CC therapeutic purposes. It is also used for diagnosis and treatment of
 CC conditions associated with aberrant Notch signalling, like cancers,
 CC including T lymphoblastic lymphoma, leukemia, adenocarcinoma, lung

CC carcinoma and cervical neoplasia. The nucleotide sequence is useful
 CC for diagnostic assays to detect expression levels of h2DX.
 XX
 SQ Sequence 19 BP; 4 A; 10 C; 4 G; 1 T; 0 other;

Query Match 100.0%; Score 12; DB 21; Length 19;
 Best Local Similarity 100.0%; Pred. No. 8e+02;
 Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GCCACCATGGCC 12
 Db 1 gccaccatggcc 12

RESULT 4
 AAS01417
 ID AAS01417 standard; DNA; 23 BP.
 XX
 AC AAS01417;
 XX
 DT 04-JUL-2001 (first entry)
 XX
 DE Mouse wild type alpha2delta-3 deletion mutant PCR primer #1.
 XX
 KW Mouse; secreted calcium channel alpha2delta subunit; alpha2delta-2;
 KW alpha2delta-3; alpha2delta-4; voltage-dependent calcium channel; VDCC;
 KW gabapentin; scintillation proximity assay; SPA; nickel flashplate assay;
 KW filter binding assay; wheat germ lectin flashplate assay; PCR primer; ss.
 XX
 OS Mus musculus.
 XX
 XX Key Location/Qualifiers
 FH misc_feature 1..8
 FT /tag= a
 FT /note= "tagged region not present in template"
 XX
 PN WO200119870-A2.
 XX
 XX 22-MAR-2001.
 XX
 PF 18-SEP-2000; 2000WO-EP09137.
 XX
 XX 16-SEP-1999; 99US-0397550.
 XX (WARN) WARNER LAMBERT CO.
 PA Brown JP, Bertelli F;
 PI WPI; 2001-235262/24.
 XX
 XX Calcium channel alpha2delta subunits, useful in e.g. SPA assays,
 PT Flashplate assays, Nickel Flashplate assays, Filter binding assays or
 PT Wheat Germ Lectin Flashplate assays -
 XX
 PS Example 12; Page 41; 160pp; English.
 XX
 XX The present sequence for deletion mutant PCR primer #1 is used with
 CC PCR primer #2 (AAS01418) to obtain mouse alpha2delta subunit deletion
 CC mutant from the wild type mouse calcium channel alpha2delta-3 subunit
 CC (AAU01026). The sequence is described in an invention relating to
 CC truncated alpha2delta-2, alpha2delta-3 or alpha2delta-4 subunit soluble
 CC proteins which retain their affinity for radioactively labelled
 CC gabapentin. The alpha2delta subunit is 1 of the components of the
 CC heteromultimeric voltage-dependent calcium channel (VDCC) complexes
 CC present in neuronal and non-neuronal tissues including heart and
 CC skeletal muscle. Numerous soluble forms of the human calcium channel
 CC alpha2delta subunits (AAU01014-AAU01024 and AAU01032-AAU01038) and 5
 CC soluble forms of the porcine calcium channel alpha2delta subunits
 CC (AAU01027-AAU01031) are described. The secreted soluble alpha2delta
 CC subunit may be used in assays e.g. scintillation proximity assay (SPA),
 CC flashplate, nickel flashplate, filter binding or wheat germ lectin
 CC flashplate assays to detect or measure the binding or interaction of a

CC ligand (e.g. gabapentin, L-Norleucine, L-Allo-Isoleucine, L-methionine,
 CC L-Leucine, L-Isoleucine, L-Valine, Spermine and/or L-Phenylalanine) of
 CC a calcium channel alpha2delta subunit.
 XX
 SQ Sequence 23 BP; 2 A; 10 C; 9 G; 2 T; 0 other;

Query Match 100.0%; Score 12; DB 22; Length 23;
 Best Local Similarity 100.0%; Pred. No. 8.1e+02;
 Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GCCACCATGGCC 12
 Db 3 gccaccatggcc 14

RESULT 5
 AAF57554
 ID AAF57554 standard; DNA; 23 BP.
 XX
 AC AAF57554;
 XX
 DT 11-JUN-2001 (first entry)
 XX
 DE Calcium channel alpha2delta subunit related primer.
 XX
 KW Calcium channel alpha2delta subunit; alpha2delta-1; cerebral cortex;
 KW nervous system disorder; pain; epilepsy; anxiety; PCR primer; ss.
 XX
 OS Synthetic.
 XX
 PN WO200120336-A2.
 XX
 PD 22-MAR-2001.
 XX
 PF 18-SEP-2000; 2000WO-EP09136.
 XX
 PR 16-SEP-1999; 99US-0397549.
 XX (WARN) WARNER LAMBERT CO.
 XX Bertelli F, Brown JP, Dissanayake V, Suman-Chauhan N, Gee NS;
 PI WPI; 2001-257902/26.
 XX
 XX Competitive binding assay for screening ligands which bind a cerebral
 PT cortical voltage-dependent calcium channel alpha2delta-1 subunit,
 PT where the ligands identified are useful for treating disorders of the
 PT nervous system, including pain -
 XX
 PS Disclosure; Page 104; 158pp; English.
 XX
 XX The invention relates to a new method for screening ligands which bind a
 CC cerebral cortical voltage-dependent calcium channel alpha2delta subunit,
 CC preferably alpha2delta-1 subunit. The method comprises contacting a
 CC secreted soluble recombinant alpha2delta-1 subunit with a ligand of
 CC interest and a labelled compound which binds the subunit, followed by
 CC measuring the level of binding of the labelled compound to alpha2delta-1
 CC subunit. The method is useful for screening ligands, preferably
 CC biologically active products that modulate a nervous system function,
 CC which bind a cerebral cortical voltage-dependent calcium channel
 CC alpha2delta-1 subunit. The ligands identified by the method are useful
 CC for treating disorders of the nervous system, including pain, epilepsy
 CC and anxiety. Sequences AAF57552-555 represent PCR primers related to
 CC calcium channel alpha2delta subunits.
 XX
 SQ Sequence 23 BP; 2 A; 10 C; 9 G; 2 T; 0 other;

Query Match 100.0%; Score 12; DB 22; Length 23;
 Best Local Similarity 100.0%; Pred. No. 8.1e+02;
 Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GCCACCATGGCC 12
 |||||||
 Db 3 gccaccatggcc 14

RESULT 6

AAQ41717
 ID AAQ41717 standard; DNA; 25 BP.

XX AC
 XX AC
 XX AC

XX 25-AUG-1993 (first entry)

XX Consensus eukaryotic ribosome binding site.

XX Vaccine; cytotoxic T lymphocyte; CTL; influenza A virus;

KW matrix protein; Ma; Pseudomonas exotoxin; cell recognition domain;

KW translocation domain; anti-viral agent; ss.

XX EP541335-A.

XX 12-MAY-1993.

XX 04-NOV-1992; 92EP-0310067.

XX 08-NOV-1991; 91US-0792507.

XX (MERI) MERCK & CO INC.

XX Donnelly JJ, Friedman A, Hawe LA, Liu MA, Marshall MS;
 PI Montgomery DL, Oliff AA, Shi X, Ulmer J;

XX WPI; 1993-154266/19.

XX Recombinant DNA encoding bacterial toxin-antigen conjugates - are
 PT useful as vaccines against viral infections, tumours and
 PT parasites

XX Example 6; Page 32; 8pp; English.

XX The construct BS-PEMI encodes a hybrid protein comprising domains I
 CC and II from Pseudomonas exotoxin fused to amino acids 2-252 of the
 CC Influenza A virus matrix protein M1. The PE-derived portion of the
 CC hybrid protein allows internalisation of the protein by an antigen-
 CC presenting cell. The hybrid protein is then processed and an antigenic
 CC segment (i.e. the Influenza A virus matrix protein) is presented on
 CC the cell surface where it elicits an immune response. BSK-PEMI was
 CC made from BS-PEMI by replacing the 21bp XhoI/HindIII fragment with
 CC a fragment (i.e. AAQ41717) encoding a consensus eukaryotic ribosome
 CC binding site. The purpose of the construct was to increase the yields
 CC of in vitro translated PEMI protein.

XX Sequence 25 BP; 6 A; 9 C; 8 G; 2 T; 0 other;

Query Match 100.0%; Score 12; DB 14; Length 25;
 Best Local Similarity 100.0%; Pred. No. 8.1e+02;
 Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GCCACCATGGCC 12
 |||||||
 Db 8 gccaccatggcc 19

RESULT 7

AAQ37109
 ID AAQ37109 standard; DNA; 25 BP.

XX AC
 XX AC

XX 20-JUL-1993 (first entry)

XX Consensus eukaryotic RBS.

XX

KW PE; Pseudomonas exotoxin; influenza A virus; M1; matrix protein;
 KW BSK-PEMI; ribosome binding site; ss.

XX Synthetic.

XX EP532090-A.

XX 17-MAR-1993.

XX 02-SEP-1992; 92EP-0202660.

XX 09-SEP-1991; 91US-0756249.

XX (MERI) MERCK & CO INC.

XX Donnelly JJ, Friedman A, Hawe LA, Liu MA, Marshall MS;
 PI Montgomery DL, Oliff AI, Shi X, Ulmer J;

XX WPI; 1993-087107/11.

XX Bacterial toxin-antigen protein conjugates - to elicit cytotoxic
 PT T-lymphocyte immune response, used for preventing viral
 PT infections, e.g. by influenza virus, HIV and human
 PT papilloma:virus

XX Claim 24; Page 78; 85pp; English.

XX Example 6 describes the construction of BSK-PEMI.

CC BSK-PEMI was made from BS-PEMI by the replacement of the 21 bp
 CC XhoI/HindIII fragment with a 24 bp (sic) fragment encoding a
 CC consensus eukaryotic ribosome binding site. The purpose of the
 CC construct was to increase the yields of in vitro translated PEMI
 CC protein.

XX Sequence 25 BP; 6 A; 9 C; 8 G; 2 T; 0 other;

Query Match 100.0%; Score 12; DB 14; Length 25;

Best Local Similarity 100.0%; Pred. No. 8.1e+02;
 Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GCCACCATGGCC 12
 |||||||
 Db 8 gccaccatggcc 19

RESULT 8

AAQ52485
 ID AAQ52485 standard; DNA; 27 BP.

XX AC
 XX AC

XX 03-JUN-1994 (first entry)

XX Primer BPI-23.

XX Polymerase chain reaction; primer; amplify; PCR; plasmid; pING4533;
 KW pING4503; bactericidal/permeability-increasing protein; BPI; IgG;
 KW heavy; chain; pMB27; fusion protein; gram negative; renal failure;
 KW bacterial infection; endotoxin related shock; metabolic acidosis;
 KW disseminated intravascular coagulation; anemia; thrombocytopenia;
 KW leukopenia; adult respiratory distress syndrome; hypotension; fever;
 KW constant region; lipopolysaccharide complement binding;
 KW placental transfer; Fc receptor binding; ss.

XX Synthetic.

XX WO9323434-A.

XX 25-NOV-1993.

XX 19-MAY-1993; 93WO-US04754.

```
XX 19-MAY-1992; 92US-0885911.
XX (XOMA ) XOMA CORP.
XX Grinna LS, Horwitz A, Theofan G;
XX WPI; 1993-386485/48.
XX
XX New fusion proteins for treating bacterial infections - comprising
XX a bactericidal-permeability-increasing protein and a immunoglobulin
XX heavy chain constant domain
XX
XX Example 1; Page 20; 75pp; English.
XX
XX The sequences given in AAQ52485-86 are primers which were used in the
XX construction of the rBPI-IgG fusion vector, pING4533. This plasmid
XX contains recombinant bactericidal/permeability-increasing protein
XX (rBPI)(1-191)ala132 with the initiating ATG in the context of the
XX consensus Kozak sequence given in AAQ52487. The rBPI fragment was fused
XX to a constant domain of an IgG heavy chain (HC). Plasmids such as
XX pING4533 may be used to produce fusion proteins which are useful for
XX the treatment of Gram negative bacterial infections and their sequelae
XX including endotoxin related shock and conditions associated with it,
XX such as disseminated intravascular coagulation, anemia,
XX thrombocytopenia, leukopenia, adult respiratory distress syndrome,
XX renal failure, hypotension, fever and metabolic acidosis. Proving
XX BPI or a fragment of it, as part of the fusion with an immuno-
XX globulin heavy chain constant region provides the potential advantages
XX of Fc receptor binding, bivalent binding to lipopolysaccharide
XX complement binding and increasing placental transfer.
XX
XX Sequence 27 BP; 5 A; 10 C; 9 G; 3 T; 0 other;

Query Match 100.0%; Score 12; DB 14; Length 27;
Best Local Similarity 100.0%; Pred. No. 8.1e+02;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GCCACCATGGCC 12
DB 10 gccaccatggcc 21

RESULT 9
AAQ67276
ID AAQ67276 standard; cDNA; 27 BP.
XX
XX AAQ67276;
XX
XX 29-MAR-1995 (first entry)
XX
XX rBPI(1-199)ala132 primer BPI-23.
XX
XX BPI; bactericidal-permeability-increasing protein; bactericide;
XX fusion protein; Gram-negative bacterium; infection; primer;
XX polymerase chain reaction; PCR; amplification; mutagenesis;
XX protein engineering; cysteine replacement analog; pING4533;
XX consensus Kozak translation sequence; rBPI(1-199)ala132; ss.
XX
XX Synthetic.
XX
XX WO9418323-A.
XX
XX 18-AUG-1994.
XX
XX 02-FEB-1994; 94WO-US01235.
XX
XX 02-FEB-1993; 93US-0013801.
XX
XX (XOMA ) XOMA CORP.
XX
XX Baltaian M, Burke D, Grinna L, Horwitz A, Theofan G;

XX 19-MAY-1992; 92US-0885911.
XX (XOMA ) XOMA CORP.
XX Grinna LS, Horwitz A, Theofan G;
XX WPI; 1993-386485/48.
XX
XX New fusion proteins for treating bacterial infections - comprising
XX a bactericidal-permeability-increasing protein and a immunoglobulin
XX heavy chain constant domain
XX
XX Example 1; Page 20; 75pp; English.
XX
XX The sequences given in AAQ52485-86 are primers which were used in the
XX construction of the rBPI-IgG fusion vector, pING4533. This plasmid
XX contains recombinant bactericidal/permeability-increasing protein
XX (rBPI)(1-191)ala132 with the initiating ATG in the context of the
XX consensus Kozak sequence given in AAQ52487. The rBPI fragment was fused
XX to a constant domain of an IgG heavy chain (HC). Plasmids such as
XX pING4533 may be used to produce fusion proteins which are useful for
XX the treatment of Gram negative bacterial infections and their sequelae
XX including endotoxin related shock and conditions associated with it,
XX such as disseminated intravascular coagulation, anemia,
XX thrombocytopenia, leukopenia, adult respiratory distress syndrome,
XX renal failure, hypotension, fever and metabolic acidosis. Proving
XX BPI or a fragment of it, as part of the fusion with an immuno-
XX globulin heavy chain constant region provides the potential advantages
XX of Fc receptor binding, bivalent binding to lipopolysaccharide
XX complement binding and increasing placental transfer.
XX
XX Sequence 27 BP; 5 A; 10 C; 9 G; 3 T; 0 other;

Query Match 100.0%; Score 12; DB 15; Length 27;
Best Local Similarity 100.0%; Pred. No. 8.1e+02;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GCCACCATGGCC 12
DB 10 gccaccatggcc 21

RESULT 10
AAQ67276
ID AAQ67276 standard; DNA; 27 BP.
XX
XX AAQ67276;
XX
XX 10-APR-1996 (first entry)
XX
XX rBPI(1-199)-Ala132 (with 27 amino acid signal) PCR primer BPI-23.
XX
XX bactericidal-permeability increasing protein; BPI; truncated; dimer;
XX endotoxic shock; heparin neutralisation; angiogenesis; inhibitor; ss.
XX
XX Synthetic.
XX
XX US5447913-A.
XX
XX 05-SEP-1995.
XX
XX 11-MAR-1994; 94US-0212132.
XX
XX 11-MAR-1994; 94US-0212132.
XX
XX (XOMA ) XOMA CORP.
XX
XX Ammons WS, Little RG;
XX
XX WPI; 1995-319904/41.
XX
XX Improved use of a bactericidal-permeability increasing protein
XX for neutralisation of heparin and treatment of endo-toxaemia, by
XX admin. as covalently linked dimer of N-terminal fragment
XX
XX Example 1; Column 15; 36pp; English.
XX
XX Compositions of N-terminal fragments (i.e. amino acids 1-193 to
XX 1-199) of human bactericidal-permeability increasing (BPI) protein
XX are used to neutralise heparin and to treat endotoxaemia. Efficacy
XX of such compns. is improved when they comprise at least 50% of the
XX truncated rBPI in dimeric form. The present sequence is that of a
XX PCR primer used in the construction of a vector coding for rBPI(1-199)
XX in which native Cys at position 132 is replaced by Ala and the
```

CC first 4 amino acids of the BPI signal peptide are removed; the
 CC codon specifying the fifth amino acid of the signal sequence (Met
 CC at position -27) is placed in the context of a consensus Kozak
 CC translation initiation sequence. The truncated rBPI is suitable
 CC for dimerisation in the presence of copper 2+ ions.
 XX
 SQ Sequence 27 BP; 5 A; 10 C; 9 G; 3 T; 0 other;

Query Match 100.0%; Score 12; DB 16; Length 27;
 Best Local Similarity 100.0%; Pred. No. 8.1e+02;
 Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GCCACCATGGCC 12
 Db 10 gccaccatggcc 21
 |||||||

RESULT 11
 AAQ97620
 ID AAQ97620 standard; cDNA; 27 BP.
 XX
 AC AAQ97620;
 XX
 DT 25-JAN-1996 (first entry)
 XX
 DE Probe/Primer for bactericidal/permeability increasing protein gene.
 XX
 KW Bactericidal permeability increasing protein; BPIP; infection;
 KW control; recombinant; ss.
 XX
 OS Synthetic.
 XX
 XX US5439807-A.
 PN
 XX 08-AUG-1995.
 PD
 XX 19-MAY-1992; 92US-0885501.
 XX
 PF 19-MAY-1993; 93US-0072063.
 PR 19-MAY-1992; 92US-0885501.
 XX
 PA (XOMA) XOMA CORP.
 XX
 PI Grinna LS;
 XX
 WPI; 1995-283094/37.
 XX
 PT Recovery of recombinant endotoxin binding protein - by culturing
 PT transformed cells in medium contg. cation exchange material, useful
 PT for treatment of bacterial infections.
 XX
 PS Disclosure; Column 21; 18pp; English.
 XX
 CC Recombinant endotoxin-binding protein comprising bactericidal/
 CC permeability increasing protein (BPIP) or its endotoxin N-terminal
 CC fragment, can be produced by growing genetically transformed host
 CC cells in a suitable medium so that BPIP is secreted into the medium.
 CC The BPIP is then isolated and can be used to control bacterial
 CC infections. The same method can be used where BPIP is replaced with
 CC a high density lipoprotein, limulus anti-LPS factor, tachyplesin or
 CC structurally related protein. This sequence is reproduced in the
 CC specification but is not referred to.
 XX
 SQ Sequence 27 BP; 5 A; 10 C; 9 G; 3 T; 0 other;

Query Match 100.0%; Score 12; DB 16; Length 27;
 Best Local Similarity 100.0%; Pred. No. 8.1e+02;
 Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GCCACCATGGCC 12
 |||||||

Db 10 gccaccatggcc 21

RESULT 12
 AAT84743
 ID AAT84743 standard; DNA; 27 BP.
 XX
 AC AAT84743;
 XX
 DT 03-NOV-1997 (first entry)
 XX
 DE Primer BPI-23 for bactericidal permeability increasing protein DNA.
 XX
 KW Primer; polymerase chain reaction; PCR; amplification; Ig; BPI;
 KW immunoglobulin; heavy chain; constant region; plasmid pMB27;
 KW preparation; hybrid; fusion; bactericidal permeability increasing;
 KW treatment; Gram-negative; bacterium; bacteria; infection;
 KW endotoxic shock; disseminated intravascular coagulation; anaemia;
 KW anaemia; thrombocytopaenia; thrombocytopenia; leukopaenia;
 KW leukopenia; adult respiratory distress syndrome; ARDS;
 KW renal failure; hypotension; fever; metabolic acidosis; ss.
 XX
 OS Synthetic.
 XX
 XX US5643570-A.
 PN
 XX 01-JUL-1997.
 PD
 XX 19-MAY-1992; 92US-0885911.
 PF
 XX 19-MAY-1993; 93US-0064693.
 PR 19-MAY-1992; 92US-0885911.
 XX
 PA (XOMA) XOMA CORP.
 XX
 PI Grinna LS, Horwitz A, Theofan G;
 XX
 WPI; 1997-350186/32.
 XX
 PT Hybrid fusion protein for treating Gram-negative bacterial
 PT infections - comprising bactericidal/permeability increasing protein
 PT and immunoglobulin heavy chain constant domain
 XX
 PS Example 1; Columns 35-36; 31pp; English.
 XX
 CC The present sequence is primer for the PCR amplification of
 CC bactericidal permeability increasing (BPI) protein DNA from the
 CC plasmid pING4533. The DNA was used in the preparation of a hybrid
 CC fusion protein, comprising an amino-terminal BPI protein and a
 CC carboxy-terminal immunoglobulin heavy chain constant region. The
 CC fusion protein can be used to treat Gram-negative bacterial
 CC infections and their sequelae, e.g. endotoxic shock, and associated
 CC conditions, e.g. disseminated intravascular coagulation, anaemia,
 CC thrombocytopaenia, leukopaenia, adult respiratory distress syndrome
 CC (ARDS), renal failure, hypotension, fever or metabolic acidosis.
 CC The fusion protein provides the potential advantages of Fc receptor
 CC binding, bivalent lipopolysaccharide binding, complement binding
 CC and increased placental transfer.
 XX
 SQ Sequence 27 BP; 5 A; 10 C; 9 G; 3 T; 0 other;

Query Match 100.0%; Score 12; DB 18; Length 27;
 Best Local Similarity 100.0%; Pred. No. 8.1e+02;
 Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GCCACCATGGCC 12
 |||||||

Db 10 gccaccatggcc 21

RESULT 13
 AAV21728

ID AAV21728 standard; cDNA; 27 BP.
 AC AAV21728;
 XX
 DT 17-AUG-1998 (first entry)
 XX
 DE P53 cDNA RT-PCR primer 1.
 XX
 KW Vector; vaccine; tumour; antigen; plasmid pITL-RHER/neu;
 KW human; p53; cancer; RT-PCR; primer; ss.
 XX
 XX Homo sapiens.
 OS Synthetic.
 XX
 PN WO9806863-A1.
 XX
 PD 19-FEB-1998.
 XX
 PF 14-AUG-1997; 97WO-US14306.
 XX
 PR 14-AUG-1996; 96US-0023931.
 XX
 PA (USSH) US DEPT HEALTH & HUMAN SERVICES.
 XX
 PI Nelson EL, Nelson PJ;
 XX
 DR WPI; 1998-159552/14.
 XX
 XX Humanised polynucleotide vectors - comprising human derived promoter
 PT and sequence acceptance site, used for the production of vaccines
 XX
 PS Example 9; Page 47; 125pp; English.
 XX
 CC RT-PCR primer 1 and RT-PCR primer 2 (see AAV21729) were used to
 CC amplify a portion of p53 cDNA (see AAV21730). The primers
 CC include mutated sequences that allow the amplified product to
 CC be incorporated into vectors such as plasmid pITL (see AAV21724).
 CC Novel humanised vectors, which can be based on pITL, comprise a
 CC human-derived promoter (or mammalian homologue) which is functional
 CC in mammalian target tissue and cells and a sequence acceptance site
 CC which accepts cDNA products from RT-PCR cloning. The vectors are
 CC non-replicating in mammalian cells but are capable of extended
 CC stable expression of the target sequence, generating an immune
 CC response in immunised individuals. The vectors selectively elicit
 CC immune responses to the target sequences with little or no immune
 CC response to the other components of the vectors. The vectors are
 CC particularly useful in accommodating monomorphic and polymorphic
 CC nucleic acid sequences encoding tumor antigens via PCR technology.
 XX
 SQ Sequence 27 BP; 3 A; 13 C; 5 G; 6 T; 0 other;

 Query Match 100.0%; Score 12; DB 19; Length 27;
 Best Local Similarity 100.0%; Pred. No. 8.1e+02;
 Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

 QY 1 GCCACCACATGGCC 12
 DB 5 gccaccatggcc 16
 |||||

 RESULT 14
 AA228010
 ID AA228010 standard; DNA; 28 BP.
 XX
 AC AA228010;
 XX
 DT 05-JAN-2000 (first entry)
 XX
 DE Human lectomedin-2 clone 2-1 amplifying primer JD#1.
 XX
 KW Human; 7-transmembrane receptor; lectomedin; lectin-binding; mucin;
 KW olfactomedin; cellular adhesion; atherosclerosis; gene therapy;

KW vascular disease; lectomedin-2; PCR primer; ss.
 XX Synthetic.
 OS Homo sapiens.
 XX
 PN WO9945111-A1.
 XX
 PD 10-SEP-1999.
 XX
 PF 04-MAR-1999; 99WO-US04676.
 XX
 PR 04-MAR-1998; 98US-0076782.
 XX
 PA (ICOS-) ICOS CORP.
 XX
 PI Hayflick JS;
 XX
 DR WPI; 1999-571596/48.
 XX
 PT New human lectomedin receptor polypeptide, used to identify specific
 PT binding partners for treating e.g. vascular disease
 XX
 PS Example 1; Page 36; 166pp; English.
 XX
 CC The invention provides purified and isolated human 7-transmembrane
 CC receptor lectomedin polypeptide or its fragments. The lectomedin
 CC polypeptide comprises extracellular lectin-binding, olfactomedin-like
 CC and mucin-like domains. The polypeptide can be produced by standard
 CC recombinant methodology. The polypeptide is involved in cellular adhesion
 CC and cytoplasmic metabolic pathways that are modulated by extracellular
 CC signaling. Specific binding to lectomedin-1 expressed on smooth muscle
 CC cells may be required for proliferation of these cells in
 CC atherosclerosis. The polypeptide is used to raise specific antibodies,
 CC and to identify specific binding agents that modulate (increase or
 CC decrease) its activity. The lectomedin nucleic acids are used as source
 CC of probes and primers, and of therapeutic antisense, ribozyme or triplex-
 CC forming agents, and in gene therapy to restore deficient lectomedin
 CC activity. Specific binding agents of lectomedin are used for treating
 CC diseases that involve lectomedin activity, e.g. vascular diseases such as
 CC atherosclerosis. The present sequence represents a primer for amplifying
 CC a human lectomedin-2 clone 2-1.
 XX
 SQ Sequence 28 BP; 6 A; 10 C; 6 G; 6 T; 0 other;

 Query Match 100.0%; Score 12; DB 20; Length 28;
 Best Local Similarity 100.0%; Pred. No. 8.1e+02;
 Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

 QY 1 GCCACCACATGGCC 12
 DB 14 gccaccatggcc 25
 |||||

 RESULT 15
 AAA96037
 ID AAA96037 standard; DNA; 28 BP.
 XX
 AC AAA96037;
 XX
 DT 26-JAN-2001 (first entry)
 XX
 DE Human lectomedin-2 cDNA 2.1 PCR primer JD#1.
 XX
 KW Human; lectomedin; seven transmembrane receptor protein;
 KW antiatherosclerotic; vasotropic; cytostatic; gene therapy;
 KW atherosclerosis; restenosis; vascular disease; peripheral neuropathy;
 KW cancer; nerve regeneration; renal cystic epithelium;
 KW uterine implantation; PCR primer; ss.
 XX
 OS Homo sapiens.
 OS Synthetic.
 XX

PN WO200052039-A2.
XX
PD 08-SEP-2000.
XX
PF 03-MAR-2000; 2000WO-US05934.
XX
PR 04-MAR-1999; 99US-0262537.
XX
PA (ICOS-) ICOS CORP.
XX
PI Hayflick JS, Fox RD, Douangpanya J, Puri KD;
XX WPI; 2000-579274/54.
DR
XX
PT Isolated human seven transmembrane receptor lectomedin polypeptide or
PT its fragment, useful for identifying agents which can treat
PT atherosclerosis, restenosis or vascular disease -
XX
PS Example 1; Page 42; 241pp; English.
XX
CC The present sequence is a PCR primer which was used in a procedure for
CC isolating human seven transmembrane receptor lectomedin polypeptides.
CC The polypeptides have a characteristic extracellular structure including
CC lectin-binding, olfactomedin-like and mucin-like domains. Modulator
CC compounds that inhibit the binding of lectomedin to a binding partner
CC (galectin-3) may be used to treat conditions such as atherosclerosis,
CC restenosis, vascular disease, peripheral neuropathy, cancer, nerve
CC regeneration, and pathologies of the renal cystic epithelia and uterine
CC implantation. Nucleotide sequences encoding the lectomedin polypeptides
CC are useful in gene therapy. The lectomedin polynucleotides and
CC polypeptides may be used to identify lectomedin binding partner
CC compounds. The polynucleotides are useful for detecting lectomedin
CC expression in cells and for identifying genetic mutations in the
CC lectomedin genes. Antisense lectomedin nucleic acids are able to inhibit
CC expression of lectomedin genes.
XX
SQ Sequence 28 BP; 6 A; 10 C; 6 G; 6 T; 0 other;

Query Match 100.0%; Score 12; DB 21; Length 28;
Best Local Similarity 100.0%; Pred. No. 8.1e+02;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 1 GCCACCATGGCC 12
| | | | | | | | | |
Db 14 gccaccatggcc 25

Search completed: January 17, 2002, 12:03:54
Job time: 17821 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: January 17, 2002, 08:54:19 ; Search time 6501.33 Seconds

(without alignments)
19.834 Million cell updates/sec

Title: US-09-242-202a-22_COPY_1_12

Perfect score: 12

Sequence: 1 GCCACCATGCC 12

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 11351937 seqs, 5372889281 residues

Total number of hits satisfying chosen parameters: 22703874

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

EST: *
1: em_estfun: *
2: em_esthum: *
3: em_estin: *
4: em_estom: *
5: em_estpi: *
6: em_estba: *
7: em_estro: *
8: em_estov: *
9: em_hic: *
10: gb_est1: *
11: gb_est2: *
12: gb_hic: *
13: gb_gss: *
14: em_gss_fun: *
15: em_gss_hum: *
16: em_gss_inv: *
17: em_gss_pin: *
18: em_gss_pro: *
19: em_gss_rod: *
20: em_gss_vrt: *
21: em_gss_other: *

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
c 1	12	100.0	49	11	BI078628
c 2	12	100.0	50	11	BF166238
c 3	12	100.0	56	11	BG242301
c 4	12	100.0	91	11	BG259484
c 5	12	100.0	99	10	BE376022
c 6	12	100.0	99	10	BE400948
c 7	12	100.0	101	11	BG006903
c 8	12	100.0	110	13	AQ978348
c 9	12	100.0	115	11	BG980184
c 10	12	100.0	117	11	BF852670
c 11	12	100.0	124	10	AA863181
c 12	12	100.0	124	11	BF092125

13	12	100.0	126	11	BG511069
14	12	100.0	127	10	AA214420
15	12	100.0	130	10	AA716853
16	12	100.0	131	11	N50245
17	12	100.0	133	11	BI152476
18	12	100.0	134	10	AU083189
19	12	100.0	135	10	AA784758
20	12	100.0	139	11	BG176372
21	12	100.0	141	11	BF533900
22	12	100.0	143	11	BG921952
23	12	100.0	143	11	BG982828
24	12	100.0	143	13	AZ225473
25	12	100.0	146	11	BG282693
26	12	100.0	148	10	BE256612
27	12	100.0	151	10	AA749848
28	12	100.0	151	10	BE366298
29	12	100.0	152	11	BF236756
30	12	100.0	157	10	AA750627
31	12	100.0	159	11	BF911674
32	12	100.0	160	10	AA279415
33	12	100.0	160	10	BE294646
34	12	100.0	161	10	BE270281
35	12	100.0	162	10	AI864485
36	12	100.0	163	10	BE266769
37	12	100.0	163	10	BE736479
38	12	100.0	164	10	AI715931
39	12	100.0	165	10	BE746616
40	12	100.0	165	13	AZ069587
41	12	100.0	166	10	BE269092
42	12	100.0	167	11	BF200800
43	12	100.0	168	11	BF752320
44	12	100.0	169	13	AZ086369
45	12	100.0	170	13	AZ314228

ALIGNMENTS

RESULT 1
BI078628/c

LOCUS

DEFINITION

BI078628

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

BI078628 49 bp mRNA EST 20-JUN-2001
602872729F1 NCI_CGAP_Mam2 Mus musculus cDNA clone IMAGE:5004750 5',
mRNA sequence.
BI078628
BI078628.1 GI:14496958
EST.
house mouse.
Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1 (bases 1 to 49)
NIH-MGC http://mgc.nci.nih.gov/
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished (1999)
Contact: Robert Strausberg, Ph.D.
Email: cgapbs@mail.nih.gov
Tissue Procurement: Gilbert Smith, Ph.D.
cDNA Library Preparation: Life Technologies, Inc.
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LLAM11043 row: i column: 07
High quality sequence stop: 49.
Location/Qualifiers
1. 49
/organism="Mus musculus"
/strain="FVB/N-3"
/db_xref="taxon:10090"
/clone="IMAGE:5004750"
/clone_lib="NCI_CGAP_Mam2"
/tissue_type="tumor, biopsy sample"

```

/dev_stage="5 months"
/lab_host="DHI0B"
/notes="mammary; Vector: pCMV-SPORT6; Site_1: SalI;
Site_2: NotI; Cloned unidirectionally. Primer: Oligo dt.
Library constructed by Life Technologies. Investigator
providing samples: Gilbert Smith, NIH"
BASE COUNT      10 a      15 c      17 g      7 t
ORIGIN

Query Match      100.0%; Score 12; DB 11; Length 49;
Best Local Similarity 100.0%; Pred. No. 2.1e+04;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GCCACCATGGCC 12
|||||
Db 34 GCCACCATGGCC 23

RESULT 2
BF166238/c
LOCUS      50 bp      mRNA      EST      30-OCT-2000
DEFINITION 601776852F1 NCI_CGAP_Lu29 Mus musculus cDNA clone IMAGE:4018606 5',
mRNA sequence.
ACCESSION  BF166238
VERSION     BF166238
KEYWORDS    EST.
SOURCE      house mouse.
ORGANISM    Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1 (bases 1 to 50)
NIH-MGC http://mgc.nci.nih.gov/.
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished (1999)
Contact: Robert Strausberg, Ph.D.
Email: cgapbs-r@mail.nih.gov
Tissue Procurement: Gilbert Smith, Ph.D.
CDNA Library Preparation: Life Technologies, Inc.
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: L1AM9269 row: g column: 23
High quality sequence stop: 50.
FEATURES             source
Location/Qualifiers
1..50
/organism="Mus musculus"
/strain="CZECH II (feral)"
/db_xref="taxon:10090"
/clone="IMAGE:4018606"
/clone_lib="NCI_CGAP_Lu29"
/tissue_type="spontaneous tumor, metastatic to mammary.
Stem cell origin."
/lab_host="DHI0B"
/notes="Organ: lung; Vector: pCMV-SPORT6; Site_1: SalI;
Site_2: NotI; Cloned unidirectionally. Primer: Oligo dt.
Library constructed by Life Technologies. Investigator
providing samples: Gilbert Smith, NIH"
BASE COUNT      10 a      15 c      18 g      7 t
ORIGIN

Query Match      100.0%; Score 12; DB 11; Length 50;
Best Local Similarity 100.0%; Pred. No. 2.1e+04;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GCCACCATGGCC 12
|||||
Db 35 GCCACCATGGCC 24

```

```

RESULT 3
BG242301/c
LOCUS      56 bp      mRNA      EST      13-FEB-2001
DEFINITION 602354472F1 NCI_CGAP_Mam1 Mus musculus cDNA clone IMAGE:4483048 5',
mRNA sequence.
ACCESSION  BG242301
VERSION     BG242301
KEYWORDS    EST.
SOURCE      house mouse.
ORGANISM    Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1 (bases 1 to 56)
NIH-MGC http://mgc.nci.nih.gov/.
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished (1999)
Contact: Robert Strausberg, Ph.D.
Email: cgapbs-r@mail.nih.gov
Tissue Procurement: Gilbert Smith, Ph.D.
CDNA Library Preparation: Life Technologies, Inc.
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: L1AM10320 row: o column: 17
High quality sequence stop: 56.
FEATURES             source
Location/Qualifiers
1..56
/organism="Mus musculus"
/strain="FVB/N"
/db_xref="taxon:10090"
/clone="IMAGE:4483048"
/clone_lib="NCI_CGAP_Mam1"
/tissue_type="tumor, biopsy sample"
/dev_stage="3 months, virgin"
/lab_host="DHI0B"
/notes="Organ: mammary; Vector: pCMV-SPORT6; Site_1: SalI;
Site_2: NotI; Cloned unidirectionally. Primer: Oligo dt.
Library constructed by Life Technologies. Investigator
providing samples: Gilbert Smith, NIH"
BASE COUNT      9 a      16 c      23 g      8 t
ORIGIN

Query Match      100.0%; Score 12; DB 11; Length 56;
Best Local Similarity 100.0%; Pred. No. 2.1e+04;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GCCACCATGGCC 12
|||||
Db 41 GCCACCATGGCC 30

RESULT 4
BG259484
LOCUS      91 bp      mRNA      EST      13-FEB-2001
DEFINITION 602378587F1 NIH_MGC_92 Homo sapiens cDNA clone IMAGE:4509501 5',
mRNA sequence.
ACCESSION  BG259484
VERSION     BG259484
KEYWORDS    EST.
SOURCE      human.
ORGANISM    Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 91)
NIH-MGC http://mgc.nci.nih.gov/.
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished (1999)
Contact: Robert Strausberg, Ph.D.
Email: cgapbs-r@mail.nih.gov
Tissue Procurement: ATCC

```

Query Match	100.0%;	Score 12;	DB 10;	Length 99;
Best Local Similarity	100.0%;	Pred. No. 2.2e+04;		

0.

[illegible]

Kingdom: Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Pooidae; Triticeae; Triticum.

TITLE International Triticeae EST Cooperative (ITEC): Production of Expressed Sequence Tags for Species of the Triticeae

JOURNAL Unpublished (2000)

COMMENT Contact: Langridge P

FEATURES

```

/clone="AWB009.G08"
/clone_lib="iTEC AWB Wheat Meiotic Stage Library"
/tissue_type="excised florets"
/dev_stage="meiotic stage no later than metaphase I"
/note="Vector: pSport 1 (life technologies cat. no. 18248-013); Site 1: SalI; Library constructed in pSport 1. Directionally cloned using the Superscript Plasmid System for cDNA synthesis and plasmid cloning. M13 Reverse sequencing primer used to obtain 5' sequence data. 1.4

```

```

BASE COUNT      22 a      39 c      22 g      16 t
ORIGIN

Query Match      100.0%; Score 12; DB 10; Length 99;
Best Local Similarity 100.0%; Pred. NO. 2.2e+04;

```

Qy	1	GCCACCATGGCC	12
Db	68	GCCACCATGGCC	79

```

RESULT 7
BG006903/c
LOCUS BG006903 101 bp mRNA EST 24-JAN-2001
DEFINITION IL5-GN0239-271100-273-g10_1 GN0239 Homo sapiens cDNA, mRNA
sequence.
ACCESSION BG006903
VERSION BG006903.1 GI:12450544
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 101)
AUTHORS Dias Neto,E., Garcia Correa,R., Verjovski-Almeida,S., Briones,M.R.,
Ngai,M.A., da Silva,W. Jr., Zago,M.A., Bordin,S., Costa,F.F.,
Goldman,G.H., Carvalho,A.F., Matsukuma,A., Baia,G.S., Simpson,D.H.,
Brunstein,A., deOliveira,P.S., Bucher,P., Jongeneel,C.V., O'Hare
,M.J., Soares,F., Brentani,R.R., Reis,L.F., de Souza,S.J. and
Simpson,A.J.
TITLE Shotgun sequencing of the human transcriptome with ORF expressed
sequence tags
JOURNAL Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)
MEDLINE 20202663
COMMENT Contact: Simpson A.J.G.
Laboratory of Cancer Genetics
Ludwig Institute for Cancer Research
Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP,
Brazil
Tel: +55-11-2704922
Fax: +55-11-2707001
Email: asimpson@ludwig.org.br
This sequence was derived from the FAPESP/LICR Human Cancer Genome
Project. This entry can be seen in the following URL
(http://www.ludwig.org.br/scripts/gethtml2.pl?tl=IL5&t2=IL5-GN0239-
271100-273-g10_1&t3=2000-11-27&t4=1)
Seq primer: puc 18 forward
High quality sequence stop: 101.
FEATURES
source
location/Qualifiers
1..101
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone_lib="GN0239"
/dev_stage="Adult"
/note="Organ: placenta_normal; Vector: puc18; Site_1: SmaI
; Site_2: SmaI; A mini-library was made by cloning
products derived from ORESTES PCR (U.S. Letters Patent
application No. 196,716 - Ludwig Institute for Cancer
Research) profiles into the puc 18 vector. Reverse
transcription of tissue mRNA and cDNA amplification were
performed under low stringency conditions."
BASE COUNT 30 a 28 c 24 g 19 t
ORIGIN
Query Match 100.0%; Score 12; DB 11; Length 101;
Best Local Similarity 100.0%; Pred. No. 2.2e+04;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 GCCACCATGGCC 12
| | | | | | | | | |
Db 81 GCCACCATGGCC 70

RESULT 8
AQ978348
LOCUS AQ978348 110 bp DNA GSS 29-JAN-2000
DEFINITION RPCI-23-321E6.TJ RPCI-23 Mus musculus genomic clone RPCI-23-321E6,
DNA sequence.
ACCESSION AQ978348
VERSION AQ978348.1 GI:6810649
KEYWORDS GSS.
SOURCE house mouse.
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE 1 (bases 1 to 110)
AUTHORS Zhao,S., Nierman,W., Feldblyum,T., Malek,J., Shatsman,S., Akinret
,B., Levins,M., McGann,S., Tsegaye,G., Geer,K., Krol,M., de Jong,P.
and Fraser,C.M.
TITLE Mouse BAC End Sequences from Library RPCI-23
JOURNAL Unpublished (1999)
COMMENT Other_GSSs: RPCI-23-321E6.TV
Contact: Shaying Zhao
Department of Eukaryotic Genomics
The Institute for Genomic Research
9712 Medical Center Dr., Rockville, MD 20850, USA
Tel: 301 838 0200
Fax: 301 838 0208
Email: szhao@tigr.org
Clones are derived from the mouse BAC library RPCI-23. For BAC
library availability, please contact Pieter de Jong
(pieter@dejong.med.buffalo.edu). Clones may be purchased from
BACPAC Resources (http://bacpac.med.buffalo.edu/orderingframe.htm)
or from Resea ch Genetics (info@resgen.com). BAC end page:
http://www.tigr.org/tdb/bac_ends/mouse/bac_end_intro.html
Plate: 321 row: E column: 6
Seq primer: SP6
Class: BAC ends.
FEATURES
source
location/Qualifiers
1..110
/organism="Mus musculus"
/strain="C57BL/6J"
/db_xref="taxon:10090"
/clone="RPCI-23-321E6"
/clone_lib="RPCI-23"
/sex="Female"
/lab_host="DH10B"
/note="Organ: Kidney/Brain; Vector: pBACe3.6; Site_1:
EcoRI; Site_2: EcoRI; Female C57BL/6J mouse kidney and/or
brain genomic DNA was isolated and partially digested
with a combination of EcoRI and EcoRI Methylase. Size
selected DNA was cloned into the pBACe3.6 vector at the
EcoRI sites. The ligation products were transformed into
DH10B electrocompetent cells (BRL Life Technologies)."
BASE COUNT 29 a 22 c 28 g 31 t
ORIGIN
Query Match 100.0%; Score 12; DB 13; Length 110;
Best Local Similarity 100.0%; Pred. No. 2.2e+04;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 GCCACCATGGCC 12
| | | | | | | | | |
Db 8 GCCACCATGGCC 19

RESULT 9
BG980184/c
LOCUS BG980184 115 bp mRNA EST 12-JUN-2001
DEFINITION PM3-CN0100-230101-002-h05 CN0100 Homo sapiens cDNA, mRNA sequence.
ACCESSION BG980184
VERSION BG980184.1 GI:14382919
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 115)
AUTHORS Dias Neto,E., Garcia Correa,R., Verjovski-Almeida,S., Briones,M.R.,
Ngai,M.A., da Silva,W. Jr., Zago,M.A., Bordin,S., Costa,F.F.,
Goldman,G.H., Carvalho,A.F., Matsukuma,A., Baia,G.S., Simpson,D.H.,
Brunstein,A., deOliveira,P.S., Bucher,P., Jongeneel,C.V., O'Hare
,M.J., Soares,F., Brentani,R.R., Reis,L.F., de Souza,S.J. and
Simpson,A.J.

```

TITLE Shotgun sequencing of the human transcriptome with ORF expressed sequence tags
JOURNAL Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)
MEDLINE 20202663
COMMENT Contact: Simpson A.J.G.
 Laboratory of Cancer Genetics
 Ludwig Institute for Cancer Research
 Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP, Brazil
 Tel: +55-11-2704922
 Fax: +55-11-2707001
 Email: asimpson@ludwig.org.br
 This sequence was derived from the FAPESP/LICR Human Cancer Genome project. This entry can be seen in the following URL
 (http://www.ludwig.org.br/scripts/gethtml2.pl?tl=PM3&t2=PM3-CN0100-230101-002-h05&t3=2001-01-23&t4=1)
 Seq primer: puc 18 forward
 High quality sequence stop: 114.

FEATURES
 source
 1. .115
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /clone_lib="CN0100"
 /dev_stage="Adult"
 /note="Organ: colon_normal; Vector: puc18; Site_1: SmaI; Site_2: SmaI; A mini-library was made by cloning products derived from ORESTES PCR (U.S. Letters Patent application No. 196,716 - Ludwig Institute for Cancer Research) profiles into the puc 18 vector. Reverse transcription of tissue mRNA and cDNA amplification were performed under low stringency conditions."
 34 a 23 c 32 g 26 t

BASE COUNT
 34 a 23 c 32 g 26 t

ORIGIN

Query Match 100.0%; Score 12; DB 11; Length 115;
 Best Local Similarity 100.0%; Pred. No. 2.2e+04;
 Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GCCACCATGCC 12
 |||||
 Db 105 GCCACCATGCC 94

RESULT 10
LOCUS BF852670 117 bp mRNA EST 16-JAN-2001
DEFINITION CM3-EN0194-151200-563-g04 EN0194 Homo sapiens cDNA, mRNA sequence.
ACCESSION BF852670
VERSION BF852670.1 GI:12340414
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE 1 (bases 1 to 117)
AUTHORS Dias Neto,E., Garcia Correa,R., Verjovski-Almeida,S., Briones,M.R., Nagai,M.A., da Silva,W. Jr., Zago,M.A., Bordin,S., Costa,F.F., Goldman,G.H., Carvalho,A.F., Matsukuma,A., Baia,G.S., Simpson,D.H., Brunstein,A., deOliveira,P.S., Bucher,P., Jongeneel,C.V., O'Hare,M.J., Soares,F., Brentani,R.R., Reis,L.F., de Souza,S.J. and Simpson,A.J.
 Shotgun sequencing of the human transcriptome with ORF expressed sequence tags

TITLE Shotgun sequencing of the human transcriptome with ORF expressed sequence tags
JOURNAL Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)
MEDLINE 20202663
COMMENT Contact: Simpson A.J.G.
 Laboratory of Cancer Genetics
 Ludwig Institute for Cancer Research
 Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP, Brazil
 Tel: +55-11-2704922
 Fax: +55-11-2707001.

Email: asimpson@ludwig.org.br
 This sequence was derived from the FAPESP/LICR Human Cancer Genome project. This entry can be seen in the following URL
 (http://www.ludwig.org.br/scripts/gethtml2.pl?tl=CM3&t2=CM3-EN0194-151200-563-g04&t3=2000-12-13&t4=1)
 Seq primer: puc 18 forward
 High quality sequence start: 25
 High quality sequence stop: 117.

FEATURES
 source
 1. .117
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /clone_lib="EN0194"
 /dev_stage="Adult"
 /note="Organ: lung_normal; Vector: puc18; Site_1: SmaI; Site_2: SmaI; A mini-library was made by cloning products derived from ORESTES PCR (U.S. Letters Patent application No. 196,716 - Ludwig Institute for Cancer Research) profiles into the puc 18 vector. Reverse transcription of tissue mRNA and cDNA amplification were performed under low stringency conditions."
 18 a 36 c 36 g 27 t

BASE COUNT
 18 a 36 c 36 g 27 t

ORIGIN
 Query Match 100.0%; Score 12; DB 11; Length 117;
 Best Local Similarity 100.0%; Pred. No. 2.2e+04;
 Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GCCACCATGCC 12
 |||||
 Db 82 GCCACCATGCC 93

RESULT 11
LOCUS AA863181/c 124 bp mRNA EST 29-APR-1998
DEFINITION Oq91f04_s1 NCI_CGAP_kid5 Homo sapiens cDNA clone IMAGE:1455679 3', similar to TR:Q94263 Q94263 SIMILAR TO THE RAS GENE FAMILY. ;, mRNA sequence.
ACCESSION AA863181
VERSION AA863181.1 GI:2955660
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE 1 (bases 1 to 124)
AUTHORS NCI-CGAP http://www.ncbi.nlm.nih.gov/hicgap.
TITLE National Cancer Institute, Cancer Genome Anatomy Project (CGAP), Tumor Gene Index
JOURNAL Unpublished (1997)
COMMENT Contact: Robert Strausberg, Ph.D.
 Email: cgapb-r@mail.nih.gov
 Tissue Procurement: Christopher Moskaluk, M.D., Ph.D., Michael R. Emmert-Buck, M.D., Ph.D.

CDNA Library Preparation: M. Bento Soares, Ph.D.
 CDNA Library Arrayed by: Greg Lennon, Ph.D.
 DNA Sequencing by: Washington University Genome Sequencing Center
 Clone distribution: NCI-CGAP clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at:
 www-bio.llnl.gov/bbrp/image/image.html

Trace considered overall poor quality
 Insert Length: 1333 Std Error: 0.00
 Seq primer: -40m13 fwd. Et from Amersham
 High quality sequence stop: 1.

FEATURES
 source
 1. .124
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /clone_lib="NCI_CGAP_kid5"

```

/tissue_type="2 pooled tumors (clear cell type)"
/lab_host="DH10B"
/Note="Organ: kidney; Vector: pT73D-Pac (Pharmacia) with
a modified polylinker; Site_1: Not I; Site_2: Eco RI; 1st
strand cDNA was primed with a Not I, oligo(dT) primer [5',
AAGTGGACAAATTCGCGCGCAATATTTTTTTTTTTTTTTT 3'],
double-stranded cDNA was ligated to Eco RI adaptors
(Pharmacia), digested with Not I and cloned into the Not I
and Eco RI sites of the modified pT73 vector. Library
went through one round of normalization. Library
constructed by Bento Soares and M. Fatima Bonaldo. "
26 a 25 c 53 g 20 t
BASE COUNT
ORIGIN

Query Match
Best Local Similarity 100.0%; Score 12; DB 10; Length 124;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GCCACCATGGCC 12
|||||
Db 25 GCCACCATGGCC 14

RESULT 12
BF092125
LOCUS
DEFINITION
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 124)
Dias Neto,E., Garcia Correa,R., Verjovski-Almeida,S., Briones,M.R.,
Nagai,M.A., da Silva,W. Jr., Zago,M.A., Bordin,S., Costa,F.F.,
Goldman,G.H., Carvalho,A.F., Matsukuma,A., Baia,G.S., Simpson,D.H.,
Brunstein,A., deOliveira,P.S., Bucher,P., Jongeneel,C.V., O'Hare
,M.J., Soares,F., Brentani,R.R., Reis,L.F., de Souza,S.J. and
Simpson,A.J.
Shotgun sequencing of the human transcriptome with ORF expressed
sequence tags
Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)
20202663
Contact: Simpson A.J.G.
Laboratory of Cancer Genetics
Ludwig Institute for Cancer Research
Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP,
Brazil
Tel: +55-11-2704922
Fax: +55-11-2707001
Email: asimpson@ludwig.org.br
This sequence was derived from the FAPESP/LICR Human Cancer Genome
project. This entry can be seen in the following URL
(http://www.ludwig.org.br/scripts/gethtml.pl?tl=st2-RC0-TN0079-150
900-026-hil&t3=2000-09-15&t4=1)
Seq primer: puc 18 forward
High quality sequence start: 31
High quality sequence stop: 124.
Location/Qualifiers
1. .124
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone_lib="TN0079"
/dev_stage="Adult"
/Note="Organ: testis_normal; Vector: puc18; Site_1: SmaI;
Site_2: SmaI; A mini-library was made by cloning products
derived from ORESTES PCR (U.S. Letters Patent application
No. 196,716 - Ludwig Institute for Cancer Research)
profiles into the pUC 18 vector. Reverse transcription of
tissue mRNA and cDNA amplification were performed under

```

```

low stringency conditions."
30 a 26 c 39 g 29 t
BASE COUNT
ORIGIN

Query Match
Best Local Similarity 100.0%; Score 12; DB 11; Length 124;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GCCACCATGGCC 12
|||||
Db 45 GCCACCATGGCC 56

RESULT 13
BG511069
LOCUS
DEFINITION
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
Glycine max
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
Rosidae; eurosids I; Fabales; Fabaceae; Papilionoideae; Phaseoleae;
Glycine.
1 (bases 1 to 126)
Shoemaker,R., Kelm,P., Vodkin,L., Erpelding,J., Coryell,V., Khanna
,A., Bolla,B., Marra,M., Hillier,L., Kucaba,T., Martin,J., Beck,C.,
Wylie,T., Underwood,K., Steptoe,M., Theising,B., Allen,M., Bowers
,Y., Person,B., Swaller,T., Gibbons,M., Pape,D., Harvey,N., Schurr
,R., Ritter,E., Kohn,S., Shin,T., Jackson,Y., Cardenas,M., McCann
,R., Waterston,R. and Wilsson,R.
Public Soybean EST Project
Unpublished (1999)
Contact: Shoemaker R/Public Soybean EST Project
Public Soybean EST Project
Washington University School of Medicine
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA
Tel: 314 286 1800
Fax: 314 286 1810
Email: est@watson.wustl.edu
This clone is available through: Genome Systems, Inc. 4633 World
Parkway Circle St. Louis, Missouri 63134 For further information
call: (800) 430-0030 or (314) 427-3222 FAX:(888) 919-3324 or (314)
427-3324 or contact: clones@genomesystems.com or
info@genomesystems.com web site: www.genomesystems.com.
Location/Qualifiers
1. .126
/organism="Glycine max"
/db_xref="taxon:3847"
/clone="GENOME SYSTEMS CLONE ID: Gm-cl072-179"
/clone_lib="Gm-cl072"
/tissue_type="seedlings induced for symptoms of SDS
(Sudden Death Syndrome) disease"
/dev_stage="2-3 weeks old"
/lab_host="DH10B"
/Note="vector: pbluescript II SK+; Site_1: EcoRI; Site_2:
XhoI; The cDNA library was constructed from mRNA isolated
from 2-3 week old seedlings that were induced for symptoms
of SDS (Sudden Death Syndrome) disease by the
translocation of culture filtrate of Fusarium solani f.
sp. glycines (Plant Cell Report 18:375-380). Cultivar PI
567374 is partially resistant to the disease SDS. Plant
tissue (expanded leaves, folded leaves, and new shoots)
were collected at 1, 6, 24, and 48 hrs. after inoculation
and their mRNA pooled equally for cDNA construction. The
library was prepared using the Stratagene pbluescript II
SK(+) library construction kit. Complementary DNA was
synthesized from mRNA using a primer consisting of a

```


poly(dT) sequence with an XhoI restriction site. EcoRI adaptors were ligated to the blunt-ended cDNA fragments followed by XhoI digestion. The cDNA insert is protected from XhoI digestion via methylation during first strand synthesis. The cDNA fragments were directionally cloned into the EcoRI-XhoI restriction site of the pBluescript vector. The ligated cDNA fragments were transformed into E.coli Electromax DH10B host cells. Plants were inoculated by Shuxian Li (Glen Hartman lab, University of Illinois). Library was constructed by Steve Clough (Lila Vodkin lab, University of Illinois).

```
BASE COUNT      28 a      34 c      34 g      30 t
ORIGIN
Query Match      100.0%; Score 12; DB 11; Length 126;
Best Local Similarity 100.0%; Pred. No. 2.2e+04;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GCCACCATGGCC 12
      |||||
Db 4 GCCACCATGGCC 15

RESULT 14
LOCUS      AA214420      127 bp      mRNA      EST      31-JAN-1997
DEFINITION zq91h04.r1 Stratagene hNT neuron (#937233) Homo sapiens cDNA clone
IMAGE:649399 5' similar to contains Alu repetitive element;; mRNA
sequence.
ACCESSION      AA214420
VERSION        AA214420.1 GI:1813037
KEYWORDS       EST.
SOURCE         human.
ORGANISM       Homo sapiens
Eukaryota; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE      1 (bases 1 to 127)
AUTHORS        Hillier,L., Lennon,G., Becker,M., Bonaldo,M.F., Chiapelli,B.,
Chisoe,S., Dietrich,N., DuBuque,T., Favello,A., Gish,W., Hawkins
,M., Hultman,M., Kucaba,T., Lacy,M., Le,M., Le,N., Mardis,E., Moore
,B., Morris,M., Parsons,J., Prange,C., Rifkin,L., Rohlfing,T.,
Schellenberg,K., Soares,M.B., Tan,F., Thierry-Mieg,J., Trevaskis,E.,
Underwood,K., Wohlmann,P., Waterston,R., Wilson,R. and Marra,M.
Generation and analysis of 280,000 human expressed sequence tags
Genome Res. 6 (9), 807-828 (1996)
9704478
TITLE          Contact: Wilson RK
JOURNAL        Washington University School of Medicine
MEDLINE        4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
COMMENT        Tel: 314 286 1800
                Fax: 314 286 1810
                Email: est@watson.wustl.edu
                This clone is available royalty-free through LLNL ; contact the
                IMAGE Consortium (info@image.llnl.gov) for further information.
                Seq primer: -28m13 rev1 ET from Amersham
                High quality sequence stop: 116.
                Location/Qualifiers
                1. 127
                /organism="Homo sapiens"
                /db_xref="GDB:5278679"
                /db_xref="taxon:9606"
                /clone="IMAGE:649399"
                /clone_lib="Stratagene hNT neuron (#937233)"
                /dev_stage="hNT neurons"
                /lab_host="SOLR (kanamycin resistant)"
                /note="Vector: pBluescript SK-; Site.1: EcoRI; Site.2:
                XhoI; Cloned unidirectionally. Primer: Oligo dT.
                Differentiated, post mitotic hNT neurons. Average insert
                size: 1.5 kb; Uni-ZAP XR vector; -5' adaptor sequence: 5'
                GAATTCGGCAGAG 3' -3' adaptor sequence: 5'
                CTCGAGTTTTTTTTTTTTTTT 3'"

BASE COUNT      18 a      38 c      31 g      36 t      4 others
```

```
ORIGIN
Query Match      100.0%; Score 12; DB 10; Length 127;
Best Local Similarity 100.0%; Pred. No. 2.2e+04;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GCCACCATGGCC 12
      |||||
Db 103 GCCACCATGGCC 114

RESULT 15
LOCUS      AA716853      130 bp      mRNA      EST      29-DEC-1997
DEFINITION vu62c06.r1 Soares_mammary_gland_NbMWG Mus musculus cDNA clone
IMAGE:1195978 5', mRNA sequence.
ACCESSION      AA716853
VERSION        AA716853.1 GI:2729127
KEYWORDS       EST.
SOURCE         house mouse.
ORGANISM       Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE      1 (bases 1 to 130)
AUTHORS        Marra,M., Hillier,L., Allen,M., Bowles,M., Dietrich,N., Dubuque,T.,
Geisel,S., Kucaba,T., Lacy,M., Le,M., Martin,J., Morris,M.,
Schellenberg,K., Steptoe,M., Tan,F., Underwood,K., Moore,B.,
Theising,B., Wylie,T., Lennon,G., Soares,B., Wilson,R. and
Waterston,R.
The WashU-HMI Mouse EST Project
Unpublished (1996)
Contact: Marra M/Mouse EST Project
WashU-HMI Mouse EST Project
Washington University School of MedicineP
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
Tel: 314 286 1800
Fax: 314 286 1810
Email: mouseest@watson.wustl.edu
This clone is available royalty-free through LLNL ; contact the
IMAGE Consortium (info@image.llnl.gov) for further information.
MGI:643074
Seq primer: -28m13 rev2 ET from Amersham
High quality sequence stop: 79.
                Location/Qualifiers
                1. 130
                /organism="Mus musculus"
                /strain="C57BL/6J"
                /db_xref="taxon:10090"
                /clone="IMAGE:1195978"
                /clone_lib="Soares_mammary_gland_NbMWG"
                /sex="male"
                /tissue_type="mammary gland"
                /dev_stage="4 weeks"
                /lab_host="DH10B"
                /note="Organ: mammary gland; Vector: pT7m3D-Pac (Pharmacia
                RI; lst strand cDNA was primed with a Not I - oligo(df)
                primer [5'
                TGTTCACCAATCTGAATGGGCGCGCCGCGAATGTTTTTTTTTTTTTTTTTTT
                T 3']; double-stranded cDNA was ligated to Eco RI
                adaptors (Pharmacia), digested with Not I and cloned into
                the Not I and Eco RI sites of the modified pT7m3 vector.
                RNA provided by Dr. Minoru Ko, Wayne State Univ. Library
                constructed and normalized by Bento Soares and M.Fatima
                Bonaldo."
                31 a      26 c      29 g      44 t

BASE COUNT      31 a      26 c      29 g      44 t
ORIGIN
Query Match      100.0%; Score 12; DB 10; Length 130;
Best Local Similarity 100.0%; Pred. No. 2.2e+04;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

Qy 1 GCCACCATGGCC 12
| | | | | | | | | |
Db 17 GCCACCATGGCC 28

Search completed: January 17, 2002, 08:54:22
Job time: 7134 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model
Run on: January 17, 2002, 11:49:54 ; Search time 10436.3 Seconds
(without alignments)
2445.411 Million cell updates/sec

Title: US-09-242-202A-27
Perfect score: 1547
Sequence: 1 GGTACCTGCCACCATGGGCGC.....CTCGCACAGCCTCTCCACACA 1547

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 1472140 seqs, 8248589755 residues

Total number of hits satisfying chosen parameters: 2944280

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

GenEmbl :

- 1: gb.ba.*
- 2: gb.htg.*
- 3: gb.in.*
- 4: gb.om.*
- 5: gb.ov.*
- 6: gb.pat.*
- 7: gb.ph.*
- 8: gb.pl.*
- 9: gb.pr.*
- 10: gb.ro.*
- 11: gb.sts.*
- 12: gb.sy.*
- 13: gb.un.*
- 14: gb.vi.*
- 15: em.ba.*
- 16: em.fun.*
- 17: em.hum.*
- 18: em.in.*
- 19: em.om.*
- 20: em.or.*
- 21: em.ov.*
- 22: em.pat.*
- 23: em.ph.*
- 24: em.pl.*
- 25: em.ro.*
- 26: em.sts.*
- 27: em.sy.*
- 28: em.un.*
- 29: em.vi.*
- 30: em.htgo_hum.*
- 31: em.htgo_inv.*
- 32: em.htgo_rod.*
- 33: em.htg_hum.*
- 34: em.htg_inv.*
- 35: em.htg_rod.*
- 36: em.htg_other.*

pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

8

Result No.	Score	Match	Length	DB	ID	Description
1	765	49.5	885	12	SYNPIAN7V	L08875 PIAN7 cloni
2	756.2	48.9	902	12	ATPIVX	V00089 Cloning vec
3	756.2	48.9	902	12	PIVIX	X14353 plasmid pi-
4	756.2	48.9	902	12	SYNPIVXV	L08918 pi-VX cloni
5	755	48.8	2932	6	I07209	I07209 Sequence 3
6	556.2	42.4	3322	12	AB009864	AB009864 Expressio
7	607.4	39.3	7252	6	AX003206	AX003206 Sequence
8	605.8	39.2	2538	12	ASPMIN1	Z50148 Artificial
9	604.4	39.1	5865	6	A60212	A60212 Sequence 8
10	604.4	39.1	5865	6	AR122288	AR122288 Sequence
11	604.4	39.1	6028	6	A60209	A60209 Sequence 5
12	604.4	39.1	6028	6	AR122285	AR122285 Sequence
13	604.4	39.1	6061	6	A60210	A60210 Sequence 6
14	604.4	39.1	6061	6	AR122286	AR122286 Sequence
15	604.4	39.1	6312	6	A60211	A60211 Sequence 7
16	604.4	39.1	6312	6	AR122287	AR122287 Sequence
17	604.2	39.1	11795	6	AX027785	AX027785 Sequence
18	604.2	39.1	13254	6	AR038307	AR038307 Sequence
19	604.2	39.1	13254	6	AR038321	AR038321 Sequence
20	604.2	39.1	13254	6	I58596	I58596 Sequence 15
21	604.2	39.1	13254	6	I58610	I58610 Sequence 17
22	603	39.0	1905	6	AR027070	AR027070 Sequence
23	603	39.0	1905	6	I86203	I86203 Sequence 9
24	603	39.0	1968	12	AF087567	AF087567 Cloning v
25	603	39.0	1969	12	AF086840	AF086840 Cloning v
26	603	39.0	1969	12	AF086841	AF086841 Cloning v
27	603	39.0	1969	12	AF086842	AF086842 Cloning v
28	603	39.0	1969	12	AF086843	AF086843 Cloning v
29	603	39.0	1969	12	AF086844	AF086844 Cloning v
30	603	39.0	1969	12	AF086845	AF086845 Cloning v
31	603	39.0	1969	12	AF086846	AF086846 Cloning v
32	603	39.0	1969	12	AF086847	AF086847 Cloning v
33	603	39.0	1969	12	AF086848	AF086848 Cloning v
34	603	39.0	1969	12	AF086849	AF086849 Cloning v
35	603	39.0	1969	12	AF086850	AF086850 Cloning v
36	603	39.0	1969	12	AF086851	AF086851 Cloning v
37	603	39.0	1969	12	AF086852	AF086852 Cloning v
38	603	39.0	1969	12	AF086853	AF086853 Cloning v
39	603	39.0	1969	12	AF087562	AF087562 Cloning v
40	603	39.0	1969	12	AF087563	AF087563 Cloning v
41	603	39.0	1969	12	AF087564	AF087564 Cloning v
42	603	39.0	1969	12	AF087565	AF087565 Cloning v
43	603	39.0	1969	12	AF087566	AF087566 Cloning v
44	603	39.0	1969	12	AF087664	AF087664 Cloning v
45	603	39.0	1969	12	AF087665	AF087665 Cloning v

ALIGNMENTS

RESULT 1	SYNPIAN7V	SYNPIAN7V	885 bp	DNA	circular	SYN	26-JUL-1993
LOCUS	PIAN7 cloning vector.						
DEFINITION	L08875						
ACCESSION	L08875.1	GI:310776					
VERSION							
KEYWORDS							
SOURCE	Synthetic construct DNA.						
ORGANISM	synthetic construct						
REFERENCE	1 (bases 1 to 885)						
AUTHORS	Gilbert,W.						
TITLE	Obtained from VecBase 3.0						
JOURNAL	Unpublished (1991)						
COMMENT	These data and their annotation were supplied to GenBank by Will Gilbert under the auspices of the GenBank Curator Program. PIAN7 - Cloning vector						
	ENTRY PIAN7						
	TITLE PIAN7 - Cloning vector						
	DATE 12-SEP-1986						
	#TYPE DNA CIRCULAR						

#sequence	16-DEC-1986
ACCESSION	Vb0066
SOURCE	artificial
REFERENCE	
#number	1
#citation	sequence information from New England Biolabs
Obtained	12-SEP-1986 from New England Biolabs
by magnetic tape	
Revised	16-DEC-1986 by F. Pfeiffer:
449/450 'AT' to 'TA' to match revised sequence of pBR322 COMMENT	
pIaN7	is thought to replace pIVX.
COMMENT	
The polylinker of pIaN7 contains additional BglII and XbaI sites within the M13mp8/pUC8 polylinker.	
KEYWORDS	
CROSSREFERENCE	
#parent	
VecBase(3):pIVX	
#parent	
VecBase(3):pBR322, VecSource(3):Poly8, GenBank(50):EcoTgy	
PARENT	
Features of pIaN7 (885 bp)	
residue source	
1- 202 202- 1 (c) Tyr-tRNA synthetic (GenBank(50):EcoTgy)	
198- 225 1- 28 part 1 of pUC8/M13mp8-polylinker	
234- 239 31- 36 part 2 of pUC8/M13mp8-polylinker	
240- 880 2522-3162 pBR322	
Conflict (cfl) and Mutations (mut): none	
FEATURE	
POLYLINKER EcoRI-SmaI-BamHI-Sali-PstI-BglII-XbaI-HindIII SELECTION	
#suppressor Supf	
SUMMARY pIaN7 #length 885 #checksum 9656.	
Location/Qualifiers	
i.. .885	
/organism="synthetic construct"	
/db_xref="taxon:32630"	
BASE COUNT 193 a 239 c 236 g 217 t	
ORIGIN	
Query Match 49.5%; Score 765; DB 12; Length 885;	
Best Local Similarity 95.9%; Pred. No. 6.3e-207;	
Matches 805; Conservative 0; Mismatches 10; Indels 24; Gaps 1;	
Qy 478 TTTCCGACTTTTCAGAGTCGTGGTGGGGAAGATTGCAACCTTCGAAAGTCATGAC 537	
Db 1 TTTCCGACTTTTCAGAGTCGTGGTGGGGAAGATTGCAACCTTCGAAAGTCATGAC 60	
Qy 538 GGCGAGATTAGATCTGCCTCTTTGGCGCTCGGGAACCCACCACCGGGTAATGCTTTT 597	
Db 61 GGCGAGATTAGATCTGCCTCTTTGGCGCTCGGGAACCCACCACCGGGTAATGCTTTT 120	
Qy 598 ACTGGCCTGCTCCTTTATCGGGAAGCGGGCGCATCATATAAATGACGCCGCTGTAA 657	
Db 121 ACTGGCCTGCTCCTTTATCGGGAAGCGGGCGCATCATATAAATGACGCCGCTGTAA 180	
Qy 658 AGTGTTACGTTGAGAAGAAATTC-----CTGCAGCCCCCGCG 693	
Db 181 AGTGTTACGTTGAGAAGAAATTCCTCCGGGATCCGTCGACCTGCAGATCTCTAGAAGCTTG 240	
Qy 694 CGTGTGCGGCTTTTCCATAGCTCCGCCCCCTGACGAGCATCACAAAATCAGCCT 753	
Db 241 CGTGTGCGGCTTTTCCATAGCTCCGCCCCCTGACGAGCATCACAAAATCAGCCT 300	
Qy 754 CAAGTCAGAGGTGGCGAAACCCGACAGGACTATAAAGATACCAGGCGTTTCCCCTGGAA 813	
Db 301 CAAGTCAGAGGTGGCGAAACCCGACAGGACTATAAAGATACCAGGCGTTTCCCCTGGAA 360	
Qy 814 GCTCCCTGCTGCTCCTCTGTTCCGACCCCTGCCGCTTACCGGATACCTGTCGCCCTTC 873	
Db 361 GCTCCCTGCTGCTCCTCTGTTCCGACCCCTGCCGCTTACCGGATACCTGTCGCCCTTC 420	
Qy 874 TCCTTCGGGAAGCGTGGCGCTTTCTCAATGCTACGCTGTAGGTAATCTCAGTTCGGTGT 933	

```
QY 531 CGATGACGGCAGATTAGAGTCTGCTCCCTTTGGCCGCTCGGAACCCACCACGGGTAA 590
|
|
|
Db 169 CGATGACGGCAGATTAGAGTCTGCTCCCTTTGGCCGCTCGGAACCCACCACGGGTAA 228
|
|
|
QY 591 TGCCTTTACTGGCCTGCTCCCTTATCGGAAGCGGGCGCATCATATCAAAATGACGGCC 650
|
|
|
Db 229 TGCCTTTACTGGCCTGCTCCCTTATCGGAAGCGGGCGCATCATATCAAAATGACGGCC 288
|
|
|
QY 651 GCTGTAAGTGTACGTTGAGAAAGATTCCTGCACGCCGCCGCTGCTGGCGTTTTTC 710
|
|
|
Db 289 GCTGTAAGTGTACGTTGAGAAAGATTCCTGCACGCCGCCGCTGCTGGCGTTTTTC 337
|
|
|
QY 711 CATAGGCTCGGCCCTCGACGAGCATCAAAAAATCGACGCTCAAGTCAGAGGTGGCGA 770
|
|
|
Db 338 CATAGGCTCGGCCCTCGACGAGCATCAAAAAATCGACGCTCAAGTCAGAGGTGGCGA 397
|
|
|
QY 771 AACCCGACAGGACTATAAGATACCAAGCGGTTCCGCCCTGGAAGCTCCCTCGCTCT 830
|
|
|
Db 398 AACCCGACAGGACTATAAGATACCAAGCGGTTCCGCCCTGGAAGCTCCCTCGCTCT 457
|
|
|
QY 831 CCGTTTCCGACCGCTCGGCTTACCGGATACCTGTCGCCCTTTCCCTTCGGGAAGCGTG 890
|
|
|
Db 458 CCGTTTCCGACCGCTCGGCTTACCGGATACCTGTCGCCCTTTCCCTTCGGGAAGCGTG 517
|
|
|
QY 891 GCGCTTCTCAATGCTCAGCTAGGTATCTCAGTTCGGGTGAGTTCGCTCCAAG 950
|
|
|
Db 518 GCGCTTCTCAATGCTCAGCTAGGTATCTCAGTTCGGGTGAGTTCGCTCCAAG 577
|
|
|
QY 951 CTGGGTGTGTGCACGAACCCCGTTACGCGGAGCGCTGCGCCCTTATCCGGTAAC 1010
|
|
|
Db 578 CTGGGTGTGTGCACGAACCCCGTTACGCGGAGCGCTGCGCCCTTATCCGGTAAC 637
|
|
|
QY 1011 CGTCTGAGTCCAAACCGGTAAACACGACTTATCGCCACTGCGCAGCCACTGGTAAC 1070
|
|
|
Db 638 CGTCTGAGTCCAAACCGGTAAACACGACTTATCGCCACTGCGCAGCCACTGGTAAC 697
|
|
|
QY 1071 AGGATTAGCAGCGAGGTATGTAGCGGTGCTACAGAGTTCCTGAAGTGTGGCTTAAC 1130
|
|
|
Db 698 AGGATTAGCAGCGAGGTATGTAGCGGTGCTACAGAGTTCCTGAAGTGTGGCTTAAC 757
|
|
|
QY 1131 TAGCGGTACACTAGAGGACAGTATTTGGTATCTGGGCTCTGCTGAAGCCAGTT-CCTTC 816
|
|
|
QY 1191 GAAAAGAGTTGGTAGCTCTTATCGGCAAAACAAACCCGCTGGTAGCGGTGTTTT 1250
|
|
|
Db 817 GAAAAGAGTTGGTAGCTCTTATCGGCAAAACAAACCCGCTGGTAGCGGTGTTTT 876
|
|
|
QY 1251 TTTGTTTGCAGCAGCATATACGC 1275
|
|
|
Db 877 TTTGTTTGCAGCAGCATATACGC 901
|
|
|
RESULT 3
PPIVX
LOCUS PPIVX 902 bp DNA SYN 07-FEB-1990
DEFINITION Plasmid pi-vx DNA.
ACCESSION X14353
VERSION X14353.1 GI:58268
KEYWORDS artificial sequence; origin of replication; plasmid; polylinker;
suppressor transfer RNA.
SOURCE
ORGANISM Plasmid synthetic construct
artificial sequence.
REFERENCE 1 (bases 1 to 902)
AUTHORS Seed,B.
TITLE Purification of genomic sequences from bacteriophage libraries by
recombination and selection in vivo
JOURNAL Nucleic Acids Res. 11 (8), 2427-2445 (1983)
MEDLINE 83220795
FEATURES
Location/Qualifiers
1..902
source
```

```
/organism="synthetic construct"
/plasmid="pi-Vx"
/db_xref="taxon:32630"
misc_feature 1..109
|
|
|
trna 110..212
|
|
|
misc_feature 213..902
|
|
|
BASE COUNT 192 a 249 c 238 g 223 t
ORIGIN
|
|
|
Query Match 48.9%; Score 756.2; DB 12; Length 902;
Best Local Similarity 98.1%; Pred. No. 2e-204;
Matches 790; Conservative 0; Mismatches 3; Indels 12; Gaps 2;
QY 471 CGAATCTTTTCGAGCTTTTGAAGTGTATGTGTGGGGGAAGGATTGAACTTCGAAGT 530
|
|
|
Db 109 CGAATCTTTTCGAGCTTTTGAAGTGTATGTGTGGGGGAAGGATTGAACTTCGAAGT 168
|
|
|
QY 531 CGATGACGGCAGATTAGAGTCTGCTCCCTTTGGCCGCTCGGAACCCACCACGGGTAA 590
|
|
|
Db 169 CGATGACGGCAGATTAGAGTCTGCTCCCTTTGGCCGCTCGGAACCCACCACGGGTAA 228
|
|
|
QY 591 TGCCTTTACTGGCCTGCTCCCTTATCGGAAGCGGGCGCATCATATCAAAATGACGGCC 650
|
|
|
Db 229 TGCCTTTACTGGCCTGCTCCCTTATCGGAAGCGGGCGCATCATATCAAAATGACGGCC 288
|
|
|
QY 651 GCTGTAAAGTGTACGTTGAGAAAGATTCCTGCACGCCGCCGCTTGTGGCGTTTTTC 710
|
|
|
Db 289 GCTGTAAAGTGTACGTTGAGAAAGATTCCTGCACGCCGCCGCTTGTGGCGTTTTTC 337
|
|
|
QY 711 CATAGGCTCGGCCCTCGACGAGCATCAAAAAATCGACGCTCAAGTCAGAGGTGGCGA 770
|
|
|
Db 338 CATAGGCTCGGCCCTCGACGAGCATCAAAAAATCGACGCTCAAGTCAGAGGTGGCGA 397
|
|
|
QY 771 AACCCGACAGGACTATAAGATACCAAGCGGTTCCGCCCTGGAAGCTCCCTCGCTCT 830
|
|
|
Db 398 AACCCGACAGGACTATAAGATACCAAGCGGTTCCGCCCTGGAAGCTCCCTCGCTCT 457
|
|
|
QY 831 CCGTTTCCGACCGCTCGGCTTACCGGATACCTGTCGCCCTTTCCCTTCGGGAAGCGTG 890
|
|
|
Db 458 CCGTTTCCGACCGCTCGGCTTACCGGATACCTGTCGCCCTTTCCCTTCGGGAAGCGTG 517
|
|
|
QY 891 GCGCTTCTCAATGCTCAGCTAGGTATCTCAGTTCGGGTGAGTTCGCTCCAAG 950
|
|
|
Db 518 GCGCTTCTCAATGCTCAGCTAGGTATCTCAGTTCGGGTGAGTTCGCTCCAAG 577
|
|
|
QY 951 CTGGGTGTGTGCACGAACCCCGTTACGCGGAGCGCTGCGCCCTTATCCGGTAAC 1010
|
|
|
Db 578 CTGGGTGTGTGCACGAACCCCGTTACGCGGAGCGCTGCGCCCTTATCCGGTAAC 637
|
|
|
QY 1011 CGTCTGAGTCCAAACCGGTAAACACGACTTATCGCCACTGCGCAGCCACTGGTAAC 1070
|
|
|
Db 638 CGTCTGAGTCCAAACCGGTAAACACGACTTATCGCCACTGCGCAGCCACTGGTAAC 697
|
|
|
QY 1071 AGGATTAGCAGCGAGGTATGTAGCGGTGCTACAGAGTTCCTGAAGTGTGGCTTAAC 1130
|
|
|
Db 698 AGGATTAGCAGCGAGGTATGTAGCGGTGCTACAGAGTTCCTGAAGTGTGGCTTAAC 757
|
|
|
QY 1131 TAGCGGTACACTAGAGGACAGTATTTGGTATCTGGGCTCTGCTGAAGCCAGTT-CCTTC 816
|
|
|
Db 758 TAGCGGTACACTAGAGGACAGTATTTGGTATCTGGGCTCTGCTGAAGCCAGTT-CCTTC 876
|
|
|
QY 1191 GAAAAGAGTTGGTAGCTCTTGTATCGGCAAAACAAACCCGCTGGTAGCGGTGTTTT 1250
|
|
|
Db 817 GAAAAGAGTTGGTAGCTCTTGTATCGGCAAAACAAACCCGCTGGTAGCGGTGTTTT 876
|
|
|
QY 1251 TTTGTTTGCAGCAGCATATACGC 1275
|
|
|
Db 877 TTTGTTTGCAGCAGCATATACGC 901
|
|
|
```

RESULT 4
 SYNPIVXV
 LOCUS SYNPIVXV 902 bp DNA circular SYN 26-JUL-1993
 DEFINITION Pi-VX cloning vector used for screening bacteriophage lambda.
 ACCESSION L08918
 VERSION L08918.1 GI:310786
 KEYWORDS
 SOURCE Synthetic construct DNA.
 ORGANISM
 REFERENCE
 AUTHORS
 TITLE
 JOURNAL
 COMMENT
 Obtained from VecBase 3.0
 Unpublished (1991)
 These data and their annotation were supplied to GenBank by Will
 Gilbert under the auspices of the GenBank Curator Program. Pi-VX -
 Cloning vector used for screening bacteriophage lambda ENTRY PIVX
 #TYPE DNA CIRCULAR
 TITLE Pi-VX - Cloning vector used for screening bacteriophage
 lambda
 DATE 17-DEC-1982
 ACCESSION V00010
 SOURCE artificial
 COLLECTION ATCC 39083
 REFERENCE
 #number 1
 #authors Seed B.
 #journal Nucl. Acids Res. (1983) 11: 2427-2445
 #title
 Purification of genomic sequences from bacteriophage libraries
 and selection in vivo.
 #comment bases 1-902
 #number 2
 #authors Maniatis T., Fritsch E.F., Sambrook J.
 #book (IN) Maniatis T., Fritsch E.F., Sambrook J.;
 MOLECULAR CLONING: 353-361;
 COLD SPRING HARBOR LABORATORY (1982).
 #comment BASES 1 TO 902
 REFERENCE 3
 #authors Backman K., Betlach M., Boyer H.W., Yanofsky S.
 #journal Cold Spring Harb Symp Quant Biol 43, 69-76 (1980)
 #title
 Genetic and Physical Studies on the replication of ColE1-type
 plasmids
 #comment
 the sequence of the PIVX origin is from pBR345 as published in
 this paper
 COMMENT
 The origin of replication is derived from pMB1 via pBR345 and
 sequence given for pBR345 differs at one position from that of
 pBR322. This
 may be a sequencing error.
 COMMENT
 The ATCC strain 39083 contains PIVX in combination with plasmid
 p3.
 COMMENT
 from GenBank
 USED FOR SCREENING BACTERIOPHAGE LAMBDA GENE LIBRARIES FOR
 SPECIFIC DNA SEQUENCES IN ESCHERICHIA COLI. PROBE SEQUENCES
 ARE INSERTED IN THE VECTOR PI-VX. CONTAINS A POLYLINKER, AND
 ORIGIN OF REPLICATION AND A TYROSINE
 AMBER-SUPPRESSOR GENE (SYNTHETIC SUPP GENE).
 KEY FROM TO DESCRIPTION
 ORGPKL 600
 ORIGIN OF REPLICATION KEYWORDS
 suppressor gene \
 polylinker \
 origin of replication
 CROSSREFERENCE
 #identical
 GenBank(50):PIVX, EMBL(11):ATPIVX
 #parent

VecBase(3):pBR322, GenBank(50):EcoTgy, VecSource(3):OriMB1
 #offspring
 VecBase(3):PIAN7, VecBase(3):mWB2344
 PARENT
 Features pf PIVX (902 bp)
 residue source
 4- 37 1- 34 pBR322
 1- 115 1- 115 PIVX-Polylinker
 111- 317 207- 1 (c) Tgt-trna synthetic (GenBank(50):EcoTgy)
 320- 901 1- 583 OriMB1
 320- 901 2522-3104 pBR322
 Conflict (cfl) and Mutations (mut):
 PIVX source
 cfl 811-12 T.C TAC 492-94 OriMB1
 cfl 811-12 T.C TAC 3013-15 pBR322
 FEATURE
 Supf
 POLYLINKER EcoRI-ClaI-HindIII-XbaI-BglII-PstI-BamHI-EcoRI
 SELECTION
 #suppressor Supf
 SUMMARY PIVX #length 902 #checksum 9265.
 Location/Qualifiers
 1..902
 /organism="synthetic construct"
 /db_xref="taxon:32630"
 BASE COUNT 192 a 250 c 237 g 223 t
 ORIGIN
 Query Match 48.9%; Score 756.2; DB 12; Length 902;
 Best Local Similarity 98.1%; Pred. No. 2e-204;
 Matches 790; Conservative 0; Mismatches 3; Indels 12; Gaps 2;
 QY 471 CGAATCTTTGGACTTTTGAAAGTGATGGTGGGGGAAAGGATTCGAACCTTCGAAGT 530
 DB 109 CGAATCTTTGGACTTTTGAAAGTGATGGTGGGGGAAAGGATTCGAACCTTCGAAGT 168
 QY 531 CGATGACGGCAGATTTACAGTCTGCCCTTTGGCCGTCGGGAACCCACCCACGGGTAA 590
 DB 169 CGATGACGGCAGATTTAGAGTCTGCCCTTTGGCCGTCGGGAACCCACCCACGGGTAA 228
 QY 591 TCGTTTTACTGGCTGCTGCCCTTATCGGGAAAGCGGGCGCATCATATCAATGACGGGCC 650
 DB 229 TCGTTTTACTGGCTGCTGCCCTTATCGGGAAAGCGGGCGCATCATATCAATGACGGGCC 288
 QY 651 GCTGTAAAGTCTTACGTTGAGAAAGATTCCTGCAGCCCGCGGTTGCTGGCGTTTTTC 710
 DB 289 GCTGTAAAGTCTTACGTTGAGAAAGATTCCTGCAGCCCGCGGTTGCTGGCGTTTTTC 337
 QY 711 CATAGGCTCCGCCCTCCCTGACGAGCATCACAAAATCGAGCTCAAGTCAGAGTGGCGA 770
 DB 338 CATAGGCTCCGCCCTCCCTGACGAGCATCACAAAATCGAGCTCAAGTCAGAGTGGCGA 397
 QY 771 AACCCGACAGACTATAAAGATACCAAGCGGTTTCCCTCGAAAGTCCCTCGTGGCGTCT 830
 DB 398 AACCCGACAGACTATAAAGATACCAAGCGGTTTCCCTCGAAAGTCCCTCGTGGCGTCT 457
 QY 831 CCTGTTCCGACCTCGCGCTTACCGGATACCTGTCGCCCTTCTCCCTTCGGGAAGCGTG 890
 DB 458 CCTGTTCCGACCTCGCGCTTACCGGATACCTGTCGCCCTTCTCCCTTCGGGAAGCGTG 517
 QY 891 GCGCTTCTCAATGCTCACGCTGTAGTATCTCAGTTCGGGTAGGTGCTTCGCTCCAAG 950
 DB 518 GCGCTTCTCATAGCTCACGCTGTAGTATCTCAGTTCGGGTAGGTGCTTCGCTCCAAG 577
 QY 951 CTGGGCTGTGTGACCAACCCCGGTTACGCCGACCGCTGCGCCCTATCCGGTAACTAT 1010
 DB 578 CTGGGCTGTGTGACCAACCCCGGTTACGCCGACCGCTGCGCCCTATCCGGTAACTAT 637
 QY 1011 CGTCTTGAGTCCAAACCCGGTAAAGACACGACTTATCGCCACTGGCAGCAGCGCTGTAAC 1070
 DB 638 CGTCTTGAGTCCAAACCCGGTAAAGACACGACTTATCGCCACTGGCAGCAGCGCTGTAAC 697

```
QY 1071 AGGATTACAGACGAGGTATGTAGCGGTGCTACAGAGTTCTTGAAGTGGTGGCCTAAC 1130
Db 698 AGGATTACAGACGAGGTATGTAGCGGTGCTACAGAGTTCTTGAAGTGGTGGCCTAAC 757
QY 1131 TACGCTACACTAGAACGACAGTATTGGTATCTCGCTCTGCTGAAAGCAGTTACCTTC 1190
Db 758 TACGCTACACTAGAACGACAGTATTGGTATCTCGCTCTGCTGAAAGCAGTT-CCTTC 816
QY 1191 GGAAAAAGTTGGTAGTCTTGATCCGGCAAAACAAACACCGCTGGTAGCGGTGGTTTT 1250
Db 817 GGAAAAAGTTGGTAGTCTTGATCCGGCAAAACAAACACCGCTGGTAGCGGTGGTTTT 876
QY 1251 TTTGTTTGAACGACGAGATTACGC 1275
Db 877 TTTGTTTGAACGACGAGATTACGC 901

RESULT 5
107209/c
LOCUS 107209 2932 bp PAT 02-DEC-1994
DEFINITION Sequence 3 from Patent EP 0330191.
ACCESSION I07209
VERSION I07209.1 GI:590060
KEYWORDS Unknown.
SOURCE Unknown.
ORGANISM Unclassified.
REFERENCE 1 (bases 1 to 2932)
AUTHORS Seed,B.D.O.M.B., Allen,J., Aruffo,A., Camerini,D., Lauffer,L.D.,
Oquendo,C.P., Simmons,D., Stamenkovic,I. and Stengelin,S.D.
TITLE Rapid immunoselection cloning method
JOURNAL Patent: EP 0330191-A2 3 30-AUG-1989;
FEATURES Location/Qualifiers
source 1..2932
/organism="unknown"
BASE COUNT 774 a 699 c 682 g 777 t
ORIGIN

Query Match 48.8%; Score 755; DB 6; Length 2932;
Best Local Similarity 97.6%; Pred. No. 4.4e-204;
Matches 788; Conservative 0; Mismatches 15; Indels 4; Gaps 2;

QY 469 ATCGAAATCTTCGGACATTTGAAAGTCATGCTGTGGGGAAGGATTCGAACCTTCGAA 528
Db 804 ACCAGATTCTTCGGACATTTGAAAGTCATGCTGTGGGGAAGGATTCGAACCTTCGAA 745
QY 529 GTCGATGACGGCAGATTTAGAGTCTGCTCCCTTTGGCGGCTCGGGAACCCACCACGGGT 588
Db 744 GTCGATGACGGCAGATTTAGAGTCTGCTCCCTTTGGCGGCTCGGGAACCCACCACGGGT 685
QY 589 AATGCTTTTACTGGCTGCTCCCTTTATCGGAAGCGGGCGCATCATATCAATGACGGG 648
Db 684 AATGCTTTTACTGGCTGCTCCCTTTATCGGAAGCGGGCGCATCATATCAATGACGGG 625
QY 649 CGCTGTTAAAGTTTACGTTGAGAAGAAATTCCTCAGCCCGCGCTGTGCGGCTTTT 708
Db 624 CGCTGTTAAAGTTTACGTTGAGAAGAA--CCGGCGGTAATTCGGCGGTGCTGGCGTTT 568
QY 709 TCCATAGGCTCCGCCCTCGACGAGCATCAAAAATCGAGCTCAAGTCAGAGGTGGC 768
Db 567 TCCATAGGCTCCGCCCTCGACGAGCATCAAAAATCGAGCTCAAGTCAGAGGTGGC 508
QY 769 GAAACCCGACAGGACTATAAGATACACAGGGCTTTCCCGCTGGAAGCTCCCTCGTGGCGT 828
Db 507 GAAACCCGACAGGACTATAAAGATACACAGGGCTTTCCCGCTGGAAGCTCCCTCGTGGCGT 448
QY 829 CTCCTGTTCCGACCTCGCGCTTACCGGATACCTGTCGCGCTTTCTCCCTTCGGGAAGCG 888
Db 447 CTCCTGTTCCGACCTCGCGCTTACCGGATACCTGTCGCGCTTTCTCCCTTCGGGAAGCG 388
QY 889 TGGCGCTTCTCAATGCTCAGCTGTAGGTATCTCAGTTTCGGTGTAGGTTCGTTCCGCTCCA 948
Db 1135 ACGAGTTCGGCTAGACGGTCTGACGACACGCAAACTGCGGAACGGTTGGGGGTGCACG 194
```

```
Db 387 TGGCGCTTTCTCATAGCTCACGCTAGTATCTCAGTTCGGTGTAGTTCGCTCTCCA 328
QY 949 AGCTGGCTGTGTGCACGAACCCCGCTTACGCCGACGCTGCCTTATTCGGTAACT 1008
Db 327 AGCTGGCTGTGTGCACGAACCCCGCTTACGCCGACGCTGCCTTATTCGGTAACT 268
QY 1009 ATCGTCTTTGAGTCCAAACCGGTAAGACACGAGCTTATCGCCACTGGCAGCAGCACTGGTA 1068
Db 267 ATCGTCTTTGAGTCCAAACCGGTAAGACACGAGCTTATCGCCACTGGCAGCAGCACTGGTA 208
QY 1069 ACAGATTAGACAGCGAGGTATGTAGCGGTGCTACAGAGTTCTTGAAGTGGTGGCTTA 1128
Db 207 ACAGATTAGACAGCGAGGTATGTAGCGGTGCTACAGAGTTCTTGAAGTGGTGGCTTA 148
QY 1129 ACTACGGCTACACTAGAGGACAGTATTGGTATCTGCGCTCTGCTGAAGCCAGTTACCT 1188
Db 147 ACTACGGCTACACTAGAGGACAGTATTGGTATCTGCGCTCTGCTGAAGCCAGTT-CCT 89
QY 1189 TCGAAAAAGAGTTGCTAGCTCTTGATCCGGCAAAACAAACCCGCTGGTAGCGGTGGTT 1248
Db 88 TCGAAAAAGAGTTGCTAGCTCTTGATCCGGCAAAACAAACCCGCTGGTAGCGGTGGTT 29
QY 1249 TTTTGTGTTGCAAGCAGCAGATTACGC 1275
Db 28 TTTTGTGTTGCAAGCAGCAGATTACGC 2

RESULT 6
AB009864
LOCUS AB009864 3392 bp DNA circular SYN 26-DEC-1997
DEFINITION Expression vector pME18S-FL3, complete sequence.
ACCESSION AB009864
VERSION AB009864.1 GI:2723416
KEYWORDS expression vectors plasmid:pME18S-FL3 DNA.
SOURCE Plasmid vectors
ORGANISM artificial sequence.
REFERENCE 1 (bases 1 to 3392)
AUTHORS Maruyama,K. and Sugano,S.
TITLE Direct Submission
JOURNAL Submitted (16-DEC-1997) to the DDBJ/EMBL/GenBank databases. Sumio
Sugano, The Institute of Medical Science, University of Tokyo,
Department of Virology; 4-6-1, Shirokanedai, Minatoku, Tokyo 108,
Japan (E-mail:sugano@ims.u-tokyo.ac.jp, Tel:81-3-5449-5286,
Fax:81-3-5449-5416)
REFERENCE 2 (bases 1 to 3392)
AUTHORS Maruyama,K. and Sugano,S.
TITLE pME18S-FL3: a versatile expression vector
JOURNAL Published Only in Database (1997) In press
FEATURES Location/Qualifiers
source 1..3392
/organism="vectors"
/plasmid="pME18S-FL3"
/db_xref="taxon:29278"
promoter 1..638
polyA_site 1229..1427
BASE COUNT 810 a 902 c 818 g 862 t
ORIGIN

Query Match 42.4%; Score 656.2; DB 12; Length 3392;
Best Local Similarity 70.8%; Pred. No. 6.6e-176;
Matches 990; Conservative 0; Mismatches 288; Indels 121; Gaps 4;

QY 15 TGGCGCGGATCTTTTATCACTGATAAGTTGGTGACATATTATGTTTATCACTGATAAAG 74
Db 828 TGGCGCGGATCTTTTATCACTGATAAGTTGGTGACATATTATGTTTATCACTGATAAAG 887
QY 75 TGTCAAGCATGACAAAGTTGACGCCGAATACAGTATCGTCCCGCCCTGGACGTGTGA 134
Db 888 TGTCAAGCATGACAAAGTTGACGCCGAATACAGTATCGTCCCGCCCTAGACCTGTGA 947
QY 135 ACGAGTTCGGCTAGACGGTCTGACGACACGCAAACTGCGGAACGGTTGGGGGTGCACG 194
```

Db 948 ACAGGTCGGGCTAGACGGTCTGACGACGCAAACTGGCGAAGCGTGGGGTTTCAGC 1007
QY 195 AGCGCGCGCTTTACTGGCACATTCAGGAACAGCGCGCCCTTAAGGCCATATAGGTGAGT 254
Db 1008 AGCGCGCGCTTTACTGGCACATTCAGGAACAGCGCGCGCTGCTGAGCGCACTGGCGGAAG 1067
QY 255 GGATGTC-----CCTGACCCAGCGGGGATGGGGAGACC 289
Db 1068 CCATGCTGGCGGAGAAATCATAGCACTTCGGTGGCGAGCGGACGACTGGCGCTCAT 1127
QY 290 TGTAGTCAGAGCCCCGGGAGCAGACAGCAATGCGCGCTCTCCCTCGCGAGTATAGTGA 349
Db 1128 TTTCTAGTGGGATGCCCGAGCTTTCAGGAGCGGCTGCTGCGCTACCGCCAGCACAATG 1187
QY 350 GTGAGTGCCTCTCCCTGCGCCCTGCGAAGTTGCGACTGCGAGTCCGCCACCGACTTGTCTCT--- 406
Db 1188 GCACCCCGTGTGCGAGCTGACGCTGCGGCGCGCTAGCTAGTCTAGAGAAAACCTCC 1247
QY 407 -----AATAAATAAGTTGCATCAATTTTGTCTGACTAGTGT 444
Db 1248 ACACCTCCCGCTGAACCTGAACATAAAATGAATGCAATTTGTTGTTAACTTGTAT 1307
QY 445 CCTCTAATAATTATAAGCTTGATATCGAATCTTTTCGACTTTTGAAGTGAATGCTGT 504
Db 1308 TGCAGCTTATAATGGTTACAAATAAAGCATAGCATCACAAAATTTCCACAAATAAAGCAT 1367
QY 505 GGGGAGGATTCGAACCTTCGAAGTCGATGAGCGGAGATTTAGAGTCTGCTCCCTTTGG 564
Db 1368 TTTTTCAGTCAATTCAGTGTGTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTT 1427
QY 565 CCCTCGGGAAACCCACCGAGTGAATGCTTTTACTGGCTGCTCCCTTATCGGAAGCG 624
Db 1428 GATCCCGGAGTACCGAGCTCGAATTAATTCCTTTCCGCTTCTCGCTCACTGACTCGT 1487
QY 625 GGGCGCATCATCAATCAATGACGCG----- 648
Db 1488 GCGCTCGGTCGTTGCGCTGCGGAGCGGTATCAGTCACTCAAAAGCGGTAATACGGTT 1547
QY 649 -----CGCTGTAAAGTGTACGTTGAGAAGAAATTCCTCGACGC 688
Db 1548 ATCCACAGAAATCAGGGATACCGCAGGAAGAAACATGTGAGCAAAAGCCAGCAAAAGC 1607
QY 689 C-----GCGCGCTGCTGCGCTTTTCCATAGGCTCCGCCGCCCTGACGA 733
Db 1608 CAGGAACCGTAAAGAGCGCGCTGCTGCGCTTTTCCATAGGCTCCGCCGCCCTGACGA 1667
QY 734 GCATCAAAATTCAGCTCAAGTCAGAGTGGGAAACCGCAGAGCTATATAAGATA 793
Db 1668 GCATCAAAATTCAGCTCAAGTCAGAGTGGGAAACCGCAGAGCTATATAAGATA 1727
QY 794 CCAGCGTTTCCCGCTGGAAGCTCCCTGCGCTCTCTGTTCCGACCTGCGCGTTAC 853
Db 1728 CCAGCGTTTCCCGCTGGAAGCTCCCTGCGCTCTCTGTTCCGACCTGCGCGTTAC 1787
QY 854 CGGATACCTGCGCGCTTTTCCCTTGGGAAGCTGGCGCTTTTCTCAATGCTCACCGTG 913
Db 1788 CGGATACCTGCGCGCTTTTCCCTTGGGAAGCTGGCGCTTTTCTCAATGCTCACCGTG 1847
QY 914 TAGGTATCTCAGTTCGGTGTAGTTCGCTCCAGCTGGGCTGTGTGACGAAACCC 973
Db 1848 TAGGTATCTCAGTTCGGTGTAGTTCGCTCCAGCTGGGCTGTGTGACGAAACCC 1907
QY 974 GCTTACGCCGACCGCTGCGCTTTATCCGTTAATCTGTTGAGTCCAAACCCGGTAAG 1033
Db 1908 GCTTACGCCGACCGCTGCGCTTTATCCGTTAATCTGTTGAGTCCAAACCCGGTAAG 1967
QY 1034 ACAGACTTATCGCACTGCGCAGCCATGTTAAGAGATTTAGCAGAGCGAGTATGT 1093
Db 1968 ACAGACTTATCGCACTGCGCAGCCATGTTAAGAGATTTAGCAGAGCGAGTATGT 2027
QY 1094 AGCGGTGCTACAGATTTCTTGAAGTGGTGGCTTAATCGGCTACACTAGAGAGCAGT 1153
|||||

Db 2028 AGCGGTGTACAGAGTTCTTGAAGTGTGGCTAACTACGCTACACTAGAAAGCAGT 2087
QY 1154 ATTTGTATCTCGCTCTGCTGAAGCCAGTTACTTCGGAAAAAGAGTTGGTAGCTCTTG 1213
Db 2088 ATTTGTATCTCGCTCTGCTGAAGCCAGTTACTTCGGAAAAAGAGTTGGTAGCTCTTG 2147
QY 1214 ATCCGGCAAAACAAACACCGCTGTGAGCGTGGTGTGTTTGTTCGAAGCAGCAGATTAC 1273
Db 2148 ATCCGGCAAAACAAACACCGCTGTGAGCGTGGTGTGTTTGTTCGAAGCAGCAGATTAC 2207
QY 1274 GCGCAGAAAAAAGGATCT 1292
Db 2208 GCGCAGAAAAAAGGATCT 2226
RESULT 7
LOCUS AX003206 7252 bp DNA PAT 24-AUG-2000
DEFINITION Sequence 57 from Patent WO9932646.
ACCESSION AX003206
VERSION AX003206.1 GI:9927068
KEYWORDS
SOURCE synthetic construct.
ORGANISM synthetic construct
REFERENCE 1 (bases 1 to 7252)
AUTHORS Carroll, M.W. and Mitrophanous, K.
TITLE Equine infectious anaemia virus (e1av) based
JOURNAL Patent: WO 9932646-A 57 01-JUL-1999;
CARROLL MILES WILLIAM (GB); MITROPHANOUS KYRIACOS (GB)
FEATURES
source
1. 7252
/organism="synthetic construct"
/db_xref="taxon:32630"
BASE COUNT 1942 a 1686 c 1751 g 1873 t
ORIGIN
Query Match 39.3%; Score 607.4; DB 6; Length 7252;
Best Local Similarity 97.5%; Pred. No. 5.5e-162;
Matches 617; Conservative 0; Mismatches 16; Indels 0; Gaps 0;
QY 660 TGTTCACCTTGAGAAAGAAATTCCTGCAGCCGCGCGCTGTGTCGCTTTTCCATAGGCTC 719
Db 5062 TATACCGGTGAGAACATGAGACTGACTCGGCGCGCTGCTGCGCTTTTCCATAGGCTC 5121
QY 720 CGCCCGCTGACGAGCATCACAAAATCGACGCTCAAGTCAGAGTGGCGAAACCGGACA 779
Db 5122 CGCCCGCTGACGAGCATCACAAAATCGACGCTCAAGTCAGAGTGGCGAAACCGGACA 5181
QY 780 GGAATAAAGATACCAAGCGTTTCCCGCTGGAAGCTCCCTGTCGCTCCTCTGTTCCG 839
Db 5182 GGAATAAAGATACCAAGCGTTTCCCGCTGGAAGCTCCCTGTCGCTCCTCTGTTCCG 5241
QY 840 ACCCTGCGCTTACCGGATACCTGTCCGCTTTTCCCTTCGGGAAGCGTGGCGTTCT 899
Db 5242 ACCCTGCGCTTACCGGATACCTGTCCGCTTTTCCCTTCGGGAAGCGTGGCGTTCT 5301
QY 900 CAATGCTCAGCTGTAGGTATCTCAGTTCGGGTAGGTGCTTCGCTCCAGCTGGGCTGT 959
Db 5302 CAATGCTCAGCTGTAGGTATCTCAGTTCGGGTAGGTGCTTCGCTCCAGCTGGGCTGT 5361
QY 960 GTGCAAGAACCCCGTTACGCCAGCCGCTGCGCTTTATCCGTTAATCTATCGTCTTGAG 1019
Db 5362 GTGCAAGAACCCCGTTACGCCAGCCGCTGCGCTTTATCCGTTAATCTATCGTCTTGAG 5421
QY 1020 TCCAAACCGGTAAAGACAGACTTATCGCACTGGCAGCAGCAGCTGGTAACAGGATAGC 1079
Db 5422 TCCAAACCGGTAAAGACAGACTTATCGCACTGGCAGCAGCAGCTGGTAACAGGATAGC 5481
QY 1080 AGAGCGAGTATGTAGGCGGTGCTACAGAGTTCTTGAAGTGGTGGCTTAACCTAGCGCTAC 1139
|||||


```
Db 5482 AGAGCCAGGTATGTAGCGGTCTACAGAGTTCTTGAAGTGGTGGCTTAACCTACGCTAC 5541
QY 1140 ACTAGAAGACAGTATTTGGTATCTCGCTCTGCTGAAGCCAGTTACCTTCGGAAAAAGA 1199
Db 5542 ACTAGAAGACAGTATTTGGTATCTCGCTCTGCTGAAGCCAGTTACCTTCGGAAAAAGA 5601
QY 1200 GTTGGTAGTCTTGATCCGGGCAACAAACCCAGCTGGTGGTAGCGGTGTTTTTTTGGTTCG 1259
Db 5602 GTTGGTAGTCTTGATCCGGGCAACAAACCCAGCTGGTGGTAGCGGTGTTTTTTTGGTTCG 5661
QY 1260 AAGCAGCAGATTACGGCGCAGAAAAAAGGATCT 1292
Db 5662 AAGCAGCAGATTACGGCGCAGAAAAAAGGATCT 5694

RESULT 8
ASPMIN1/c
LOCUS ASPMIN1 2538 bp DNA SYN 22-APR-1996
DEFINITION Artificial DNA for plasmid pMin1.
ACCESSION 250148
VERSION 250148.1 GI:1052554
KEYWORDS plasmid; tetracyclin-resistance protein.
SOURCE synthetic construct.
ORGANISM synthetic construct.
artificial sequence.
REFERENCE 1 (bases 1 to 2538)
AUTHORS Haas,R.
TITLE Improved TnMax mini-transposon system suitable for sequencing,
shuttle mutagenesis and gene fusions
JOURNAL Gene (1995) In press
AUTHORS Haas,R.
TITLE Direct Submission
JOURNAL Submitted (21-JUL-1995) Rainer Haas, Infektionsbiologie,
Max-Planck-Institut fuer, Biologie, Spemannstr. 34, Tuebingen,
B.W., D-72076, Germany
REFERENCE 3 (bases 1 to 2538)
AUTHORS Kahrs,A.F., Odenbret,S., Schmitt,W., Heuermann,D., Meyer,T.F. and
Haas,R.
TITLE An improved TnMax mini-transposon system suitable for sequencing,
shuttle mutagenesis and gene fusions
JOURNAL Gene. 167 (1-2), 53-57 (1995)
MEDLINE 96144248
PUBMED 8566811
FEATURES
source
Location/Qualifiers
1..2538
/organism="synthetic construct"
/db_xref="taxon:32630"
/lab_host="Escherichia coli K12"
1..45
/note="multiple cloning site (mcs)"
67..100
/evidence=experimental
186..1376
/codon_start=1
/transl_table=11
/product="pMin1, tetracyclin-resistance protein"
/protein_id="CAA90509.1"
/db_xref="GI:1052555"
/translation="MKSNNALIVLGTVTLDAVGIGLVNPLPGLLRDIVHSDTASH
YGVLLALYALMQLCAPVLGALSDFRRPRLVLLASLLGATIDYATMATTPVLWILYAG
RIVAGITGATGAVAGAYTADITDGEDRARHFLMSACFGVGNVAGPVAGGLIATSLH
APFLAAAVLNGINLLGCLFMQESHKGRPPMLRAFNPVSVSFRWARGMTVAALMTV
FFIMLVGQPAALVWIFCEDRFWSATMIGLSLAVFGILHALAQAFVTGPATKRFGE
KOATLIGAAADLVLLAFATRGWMAFPIMLLLASGIGMPALQAMLSROVDDHOG
QLQGSALATSLTSGTLIVTAYASASTWNLAWIVGAALYVLCVLPALRRGAWSR
Arst"
1392..1766
/note="orit from plasmid RP4"
1767..2507
/note="replication origin of plasmid pBR322 (oriColEI)"
2508..2537
/note="bacteriophage fd terminator (tfd)"

rep_origin
rep_origin
terminator
```

```
BASE COUNT 508 a 717 c 708 g 605 t
ORIGIN
Query Match 39.2%; Score 605.8; DB 12; Length 2538;
Best Local Similarity 97.3%; Pred. No. 1.6e-161;
Matches 616; Conservative 0; Mismatches 17; Indels 0; Gaps 0;
QY 690 GCGCGCTTCTCGCGCTTTTTCATAGGCTCCGCCCTTGACGAGCATCACAAAAATCGA 749
Db 2388 GCGCGCTTCTCGCGCTTTTTCATAGGCTCCGCCCTTGACGAGCATCACAAAAATCGA 2329
QY 750 CGCTCAAGTCAGAGTGGCGAAACCCGACGAGGACTATAAAGATACACAGGCTTCCCCCT 809
Db 2328 CGCTCAAGTCAGAGTGGCGAAACCCGACGAGGACTATAAAGATACACAGGCTTCCCCCT 2269
QY 810 GGAAGCTCCCTCGTGGCTCTCTCTTCCGACCTCCCGCTTACCGGATACCTGTCCGCC 869
Db 2268 GGAAGCTCCCTCGTGGCTCTCTCTTCCGACCTCCCGCTTACCGGATACCTGTCCGCC 2209
QY 870 TTTCTCCCTTCGGGAAGCGTGGCGCTTCTCAATGCTCAGCTGTAGGTATCTCAGTTCCG 929
Db 2208 TTTCTCCCTTCGGGAAGCGTGGCGCTTCTCAATGCTCAGCTGTAGGTATCTCAGTTCCG 2149
QY 930 GTGTAGTCTGTGCTCCAAAGCTGGGCTGTGTGCACGAACCCCGCTTCAGCCCGACCGC 989
Db 2148 GTGTAGTCTGTGCTCCAAAGCTGGGCTGTGTGCACGAACCCCGCTTCAGCCCGACCGC 2089
QY 990 TCGCGCTTATCGGTAACATATCGTCTTGAAGTCCAAACCCCGGTAGACACGACTTATCGCCA 1049
Db 2088 TCGCGCTTATCGGTAACATATCGTCTTGAAGTCCAAACCCCGGTAGACACGACTTATCGCCA 2029
QY 1050 CTGGCAGCAGCCACTGCTAACAGGATTAGCAGAGCAGGATGTAGCGGTGCTACAGAG 1109
Db 2028 CTGGCAGCAGCCACTGCTAACAGGATTAGCAGAGCAGGATGTAGCGGTGCTACAGAG 1969
QY 1110 TTCTTGAAGTGTGGCTTAACCTACGCTACACTAGAGGACAGTATTTGGTATCTCGCT 1169
Db 1968 TTCTTGAAGTGTGGCTTAACCTACGCTACACTAGAGGACAGTATTTGGTATCTCGCT 1909
QY 1170 CTGCTGAAGCCAGTTACCTTCGGAAAAAGAGTTGGTAGCTCTTGATCCGCAACCAACC 1229
Db 1908 CTGCTGAAGCCAGTTACCTTCGGAAAAAGAGTTGGTAGCTCTTGATCCGCAACCAACC 1849
QY 1230 ACCGCTGTAGCGGTGGTTTTTTTGTTCGACGACAGATTACGCGCAAGAAAAAAGGA 1289
Db 1848 ACCGCTGTAGCGGTGGTTTTTTTGTTCGACGACAGATTACGCGCAAGAAAAAAGGA 1789
QY 1290 TCTGGGGATCCGGAGAGCTCACTCTAGATGAG 1322
Db 1788 TCTCAAGAAGATCCTTTGATCTCTGCAGATCAG 1756

RESULT 9
A60212
LOCUS A60212 5865 bp DNA circular PAT 06-MAR-1998
DEFINITION Sequence 8 from Patent WO9708330.
ACCESSION A60212
VERSION A60212.1 GI:3715220
KEYWORDS unidentified.
SOURCE unidentified.
ORGANISM unclassified.
REFERENCE 1 (bases 1 to 5865)
AUTHORS Collins,M.K., Weiss,R.A., Takeuchi,Y. and Cosset,F.
TITLE EXPRESSION SYSTEMS
JOURNAL Patent: WO 9708330-A 8 06-MAR-1997;
CANCER RES CAMPAIGN TECH (GB)
FEATURES
Location/Qualifiers
1..5865
/organism="unidentified"
/db_xref="taxon:32644"
BASE COUNT 1492 a 1602 c 1390 g 1373 t 8 others
```

ORIGIN

Query Match 39.1%; Score 604.4; DB 6; Length 5865;
Best Local Similarity 91.4%; Pred. No. 3.9e-161;
Matches 638; Conservative 0; Mismatches 60; Indels 0; Gaps 0;

Qy 595 TTTACTGCCCTGCTCCCTTATCGGGAACGGGGGCATCATATCAAAATGACGCGCGCTG 654
Db 3752 TTATCATGCTCGATCCAGATCTGGCCCATCGCGCCGCGATCGATNNNACATGTGAG 3811

Qy 655 TAAAGTGTACGTTGAGAAAGAAATTCCTGACGCGCGCTGCTGGCTGCTGCTTTTCCATA 714
Db 3812 CAAAAGGCCACAAAGGCCAGAACCGCTAAAGAGCGCGCTGCTGGCTGCTTTTCCATA 3871

Qy 715 GGCTCCGCCCCCTGACGAGCATCACAATAATCGACGCTCAAGTCAGAGGTGGGGAAC 774
Db 3872 GGCTCCGCCCCCTGACGAGCATCACAATAATCGACGCTCAAGTCAGAGGTGGGGAAC 3931

Qy 775 CGACAGGACTATAAGATACCAAGGCTTTCCCTGGAAGCTCCCTCGTGGCTCTCCTG 834
Db 3932 CGACAGGACTATAAGATACCAAGGCTTTCCCTGGAAGCTCCCTCGTGGCTCTCCTG 3991

Qy 835 TTCGACCTCGCGCTTACCGGATACCTGTCGCGCTTTCTCCCTTCGGGAAGCGTGGCG 894
Db 3992 TTCGACCTCGCGCTTACCGGATACCTGTCGCGCTTTCTCCCTTCGGGAAGCGTGGCG 4051

Qy 895 TTTCTCAATGCTACGCTGTAGGTATCTCAGTTCGGTGTAGGTGCTTCGCTCCAAGCTGG 954
Db 4052 TTTCTCAATGCTACGCTGTAGGTATCTCAGTTCGGTGTAGGTGCTTCGCTCCAAGCTGG 4111

Qy 955 GCTGTGTCACGAACCCCGCTTCAGCCCGACCGCTGCGCTTTATCCGGTAACTATCGTC 1014
Db 4112 GCTGTGTCACGAACCCCGCTTCAGCCCGACCGCTGCGCTTTATCCGGTAACTATCGTC 4171

Qy 1015 TTGAGTCCACCCCGGTAGACACGACTATCCGACCTGGGAGAGCCACTGGTAACAGGA 1074
Db 4172 TTGAGTCCACCCCGGTAGACACGACTATCCGACCTGGGAGAGCCACTGGTAACAGGA 4231

Qy 1075 TTAGCAGGCGAGGTATGAGCGGTGCTACAGAGTTCTTCAAGTGGTGGCTAACTACG 1134
Db 4232 TTAGCAGGCGAGGTATGAGCGGTGCTACAGAGTTCTTCAAGTGGTGGCTAACTACG 4291

Qy 1135 GCTACACTAGAAGACAGTATTTGGTATCTGCGCTCTGCTGAAGCCAGTTACCTTCGGA 1194
Db 4292 GCTACACTAGAAGACAGTATTTGGTATCTGCGCTCTGCTGAAGCCAGTTACCTTCGGA 4351

Qy 1195 AAAGAGTTGGTAGCTCTTGATCCGGCAACAAACACCGCTGGTAGCGGTGTTTTTTT 1254
Db 4352 AAAGAGTTGGTAGCTCTTGATCCGGCAACAAACACCGCTGGTAGCGGTGTTTTTTT 4411

Qy 1255 TTTGCAAGCAGCAGATTACGCGCAAGAAAAGGATCT 1292
Db 4412 TTTGCAAGCAGCAGATTACGCGCAAGAAAAGGATCT 4449

RESULT 10
A60209
LOCUS A60209 5865 bp DNA PAT 16-MAY-2001
DEFINITION Sequence 8 from patent US 6165715.
ACCESSION AR122288
VERSION AR122288.1 GI:14106605
KEYWORDS Unknown.
SOURCE Unclassified.
ORGANISM Unclassified.
REFERENCE 1 (bases 1 to 5865)
AUTHORS Collins,M.K., KatherineLevinge, Weiss, R. Anthony, Takeuchi,Y. and Cosset,F.
TITLE Expression systems
JOURNAL Patent: US 6165715-A 8 26-DEC-2000;
FEATURES Location/Qualifiers
source l. 5865

BASE COUNT 1492 a 1602 c 1390 g 1373 t 8 others
ORIGIN

Query Match 39.1%; Score 604.4; DB 6; Length 5865;
Best Local Similarity 91.4%; Pred. No. 3.9e-161;
Matches 638; Conservative 0; Mismatches 60; Indels 0; Gaps 0;

Qy 595 TTTACTGCCCTGCTCCCTTATCGGGAACGGGGGCATCATATCAAAATGACGCGCGCTG 654
Db 3752 TTATCATGCTCGATCCAGATCTGGCCCATCGCGCCGCGATCGATNNNACATGTGAG 3811

Qy 655 TAAAGTGTACGTTGAGAAAGAAATTCCTGACGCGCGCTGCTGGCTGCTGCTTTTCCATA 714
Db 3812 CAAAAGGCCACAAAGGCCAGAACCGCTAAAGAGCGCGCTGCTGGCTGCTTTTCCATA 3871

Qy 715 GGCTCCGCCCCCTGACGAGCATCACAATAATCGACGCTCAAGTCAGAGGTGGGGAAC 774
Db 3872 GGCTCCGCCCCCTGACGAGCATCACAATAATCGACGCTCAAGTCAGAGGTGGGGAAC 3931

Qy 775 CGACAGGACTATAAGATACCAAGGCTTTCCCTGGAAGCTCCCTCGTGGCTCTCCTG 834
Db 3932 CGACAGGACTATAAGATACCAAGGCTTTCCCTGGAAGCTCCCTCGTGGCTCTCCTG 3991

Qy 835 TTCGACCTCGCGCTTACCGGATACCTGTCGCGCTTTCTCCCTTCGGGAAGCGTGGCG 894
Db 3992 TTCGACCTCGCGCTTACCGGATACCTGTCGCGCTTTCTCCCTTCGGGAAGCGTGGCG 4051

Qy 895 TTTCTCAATGCTACGCTGTAGGTATCTCAGTTCGGTGTAGGTGCTTCGCTCCAAGCTGG 954
Db 4052 TTTCTCAATGCTACGCTGTAGGTATCTCAGTTCGGTGTAGGTGCTTCGCTCCAAGCTGG 4111

Qy 955 GCTGTGTCACGAACCCCGCTTCAGCCCGACCGCTGCGCTTTATCCGGTAACTATCGTC 1014
Db 4112 GCTGTGTCACGAACCCCGCTTCAGCCCGACCGCTGCGCTTTATCCGGTAACTATCGTC 4171

Qy 1015 TTGAGTCCACCCCGGTAGACACGACTATCCGACCTGGGAGAGCCACTGGTAACAGGA 1074
Db 4172 TTGAGTCCACCCCGGTAGACACGACTATCCGACCTGGGAGAGCCACTGGTAACAGGA 4231

Qy 1075 TTAGCAGGCGAGGTATGAGCGGTGCTACAGAGTTCTTCAAGTGGTGGCTAACTACG 1134
Db 4232 TTAGCAGGCGAGGTATGAGCGGTGCTACAGAGTTCTTCAAGTGGTGGCTAACTACG 4291

Qy 1135 GCTACACTAGAAGACAGTATTTGGTATCTGCGCTCTGCTGAAGCCAGTTACCTTCGGA 1194
Db 4292 GCTACACTAGAAGACAGTATTTGGTATCTGCGCTCTGCTGAAGCCAGTTACCTTCGGA 4351

Qy 1195 AAAGAGTTGGTAGCTCTTGATCCGGCAACAAACACCGCTGGTAGCGGTGTTTTTTT 1254
Db 4352 AAAGAGTTGGTAGCTCTTGATCCGGCAACAAACACCGCTGGTAGCGGTGTTTTTTT 4411

Qy 1255 TTTGCAAGCAGCAGATTACGCGCAAGAAAAGGATCT 1292
Db 4412 TTTGCAAGCAGCAGATTACGCGCAAGAAAAGGATCT 4449

RESULT 11
A60209
LOCUS A60209 6028 bp DNA circular PAT 06-MAR-1998
DEFINITION Sequence 5 from Patent WO9708330.
ACCESSION A60209
VERSION A60209.1 GI:3715217
KEYWORDS unidentified.
SOURCE unidentified.
ORGANISM unidentified.
REFERENCE 1 (bases 1 to 6028)
AUTHORS Collins,M.K., Weiss,R.A., Takeuchi,Y. and Cosset,F.
TITLE EXPRESSION SYSTEMS
JOURNAL Patent: WO 9708330-A 5 06-MAR-1997;
CANCER RES CAMPAIGN TECH (GB)

FEATURES
source
Location/Qualifiers
1..6028
/organism="unidentified"
/db_xref="taxon:32644"
BASE COUNT 1515 a 1637 c 1499 g 1369 t 8 others
ORIGIN

Query Match 39.1%; Score 604.4; DB 6; Length 6028;
Best Local Similarity 91.4%; Pred. No. 3.9e-161;
Matches 638; Conservative 0; Mismatches 60; Indels 0; Gaps 0;

QY 595 TTTACTGGCTGTCCCTTATCGGGAAGGGCGGCATCATATCAAAATGACGCGCGCTG 654
DB 3915 TTATCATGTCTGGATCCAGATCTGGGCCCATGCGGCCGGATCGATNNNACATGTGAG 3974
QY 655 TAAAGTGTACGTTGAGAAGAAATTCCTGCAGCCCGCCGCTTCTGCTGGCGTTTTCCTATA 714
DB 3975 CAAAAGGCCAGCAAAAGGCCAGCAACCGTAAAGGCCGCTTCTGCTGGCGTTTTCCTATA 4034
QY 715 GGCTCGCCCGCCCTGACGAGCATCAGCAAAATCGAGCTCAAGTCAGAGGTGCGCAAAACC 774
DB 4035 GGCTCGCCCGCCCTGACGAGCATCAGCAAAATCGAGCTCAAGTCAGAGGTGCGCAAAACC 4094
QY 775 CGACAGGACTATAAAGATACCAAGCGCTTTCCCGCTGGAAGCTCCCTCGCGCTCTCCCTG 834
DB 4095 CGACAGGACTATAAAGATACCAAGCGCTTTCCCGCTGGAAGCTCCCTCGCTCTCCCTG 4154
QY 835 TTCCGACCTCGCGCTTACCGGATACCTGCGCCCTTCTCCCTTTCGCGGAAGCGTGGCGC 894
DB 4155 TTCCGACCTCGCGCTTACCGGATACCTGCGCCCTTCTCCCTTTCGCGGAAGCGTGGCGC 4214
QY 895 TTTCTCAATGCTCACGCTGTAGTATCTCAGTTCGGTGTAGGTCGTTCCGCTCAAGCTGG 954
DB 4215 TTTCTCAATGCTCACGCTGTAGTATCTCAGTTCGGTGTAGGTCGTTCCGCTCAAGCTGG 4274
QY 955 GCTGTGTGACGAACCCCGCTTACGCGGACCGCTTATCGGCTTATCCGCTTAACCTACGTC 1014
DB 4275 GCTGTGTGACGAACCCCGCTTACGCGGACCGCTTATCGGCTTATCCGCTTATCCGCT 4334
QY 1015 TTGAGTCCAAACCCGGTAAAGACAGCTTATCGCCACTGGCAGCAGCCACTGGTAAACGGA 1074
DB 4335 TTGAGTCCAAACCCGGTAAAGACAGCTTATCGCCACTGGCAGCAGCCACTGGTAAACGGA 4394
QY 1075 TTAGCAGAGGAGGTATGAGCGGTGCTACAGATTCCTGAAGTGGTGGCCTTAACCTACG 1134
DB 4395 TTAGCAGAGGAGGTATGAGCGGTGCTACAGATTCCTGAAGTGGTGGCCTTAACCTACG 4454
QY 1135 GCTACACTAGAAGGACAGATTTGGTATCTGCGCTCTGCTGAAGCCAGTTACCTTCGGAA 1194
DB 4455 GCTACACTAGAAGGACAGATTTGGTATCTGCGCTCTGCTGAAGCCAGTTACCTTCGGAA 4514
QY 1195 AAAGAGTTGGTAGCTTGTATCCGGCAAAACCAACCCGCTGTAGCGGTGGTTTTTTTG 1254
DB 4515 AAAGAGTTGGTAGCTTGTATCCGGCAAAACCAACCCGCTGTAGCGGTGGTTTTTTTG 4574
QY 1255 TTGCAAGCAGCAGATTACGCGCAGAAAAAAGGATCT 1292
DB 4575 TTGCAAGCAGCAGATTACGCGCAGAAAAAAGGATCT 4612

RESULT 12
LOCUS AR122285 6028 bp DNA PAT 16-MAY-2001
DEFINITION Sequence 5 from patent US 6165715.
ACCESSION AR122285
VERSION AR122285.1 GI:14106602
KEYWORDS
SOURCE Unknown.
ORGANISM Unclassified.
REFERENCE 1 (bases 1 to 6028)
AUTHORS Collins,M.KatherineLevinge, Weiss,R.Anthony, Takeuchi,Y. and

Cosset,F.
Expression systems
Patent: US 6165715-A 5 26-DEC-2000;
Location/Qualifiers
1..6028
/organism="unknown"
BASE COUNT 1515 a 1637 c 1499 g 1369 t 8 others
ORIGIN

Query Match 39.1%; Score 604.4; DB 6; Length 6028;
Best Local Similarity 91.4%; Pred. No. 3.9e-161;
Matches 638; Conservative 0; Mismatches 60; Indels 0; Gaps 0;

QY 595 TTTACTGGCTGTCCCTTATCGGGAAGGGCGGCATCATATCAAAATGACGCGCGCTG 654
DB 3915 TTATCATGTCTGGATCCAGATCTGGGCCCATGCGGCCGGATCGATNNNACATGTGAG 3974
QY 655 TAAAGTGTACGTTGAGAAGAAATTCCTGCAGCCCGCCGCTTCTGCTGGCGTTTTCCTATA 714
DB 3975 CAAAAGGCCAGCAAAAGGCCAGCAACCGTAAAGGCCGCTTCTGCTGGCGTTTTCCTATA 4034
QY 715 GGCTCGCCCGCCCTGACGAGCATCAGCAAAATCGAGCTCAAGTCAGAGGTGCGCAAAACC 774
DB 4035 GGCTCGCCCGCCCTGACGAGCATCAGCAAAATCGAGCTCAAGTCAGAGGTGCGCAAAACC 4094
QY 775 CGACAGGACTATAAAGATACCAAGCGCTTTCCCGCTGGAAGCTCCCTCGCTCTCCCTG 834
DB 4095 CGACAGGACTATAAAGATACCAAGCGCTTTCCCGCTGGAAGCTCCCTCGCTCTCCCTG 4154
QY 835 TTCCGACCTCGCGCTTACCGGATACCTGCGCCCTTCTCCCTTTCGCGGAAGCGTGGCGC 894
DB 4155 TTCCGACCTCGCGCTTACCGGATACCTGCGCCCTTCTCCCTTTCGCGGAAGCGTGGCGC 4214
QY 895 TTTCTCAATGCTCACGCTGTAGTATCTCAGTTCGGTGTAGGTCGTTCCGCTCAAGCTGG 954
DB 4215 TTTCTCAATGCTCACGCTGTAGTATCTCAGTTCGGTGTAGGTCGTTCCGCTCAAGCTGG 4274
QY 955 GCTGTGTGACGAACCCCGCTTACGCGGACCGCTTATCGGCTTATCCGCTTAACCTACGTC 1014
DB 4275 GCTGTGTGACGAACCCCGCTTACGCGGACCGCTTATCGGCTTATCCGCTTATCCGCT 4334
QY 1015 TTGAGTCCAAACCCGGTAAAGACAGCTTATCGCCACTGGCAGCAGCCACTGGTAAACGGA 1074
DB 4335 TTGAGTCCAAACCCGGTAAAGACAGCTTATCGCCACTGGCAGCAGCCACTGGTAAACGGA 4394
QY 1075 TTAGCAGAGGAGGTATGAGCGGTGCTACAGATTCCTGAAGTGGTGGCCTTAACCTACG 1134
DB 4395 TTAGCAGAGGAGGTATGAGCGGTGCTACAGATTCCTGAAGTGGTGGCCTTAACCTACG 4454
QY 1135 GCTACACTAGAAGGACAGATTTGGTATCTGCGCTCTGCTGAAGCCAGTTACCTTCGGAA 1194
DB 4455 GCTACACTAGAAGGACAGATTTGGTATCTGCGCTCTGCTGAAGCCAGTTACCTTCGGAA 4514
QY 1195 AAAGAGTTGGTAGCTTGTATCCGGCAAAACCAACCCGCTGTAGCGGTGGTTTTTTTG 1254
DB 4515 AAAGAGTTGGTAGCTTGTATCCGGCAAAACCAACCCGCTGTAGCGGTGGTTTTTTTG 4574
QY 1255 TTGCAAGCAGCAGATTACGCGCAGAAAAAAGGATCT 1292
DB 4575 TTGCAAGCAGCAGATTACGCGCAGAAAAAAGGATCT 4612

RESULT 13
LOCUS A60210 6061 bp DNA circular PAT 06-MAR-1998
DEFINITION Sequence 6 from Patent W09708330.
ACCESSION A60210
VERSION A60210.1 GI:3715218
KEYWORDS
SOURCE unidentified.
ORGANISM unidentified.
unclassified.

REFERENCE 1 (bases 1 to 6061)
AUTHORS Collins,M.K., Weiss,R.A., Takeuchi,Y. and Cosset,F.
TITLE EXPRESSION SYSTEMS
JOURNAL PATENT: WO 9708330-A 6 06-MAR-1997;
CANCER RES CAMPAIGN TECH (GB)
FEATURES Location/Qualifiers
source 1..6061
/organism="unidentified"
/db_xref="taxon:32644"
BASE COUNT 1498 a 1665 c 1485 g 1405 t 8 others
ORIGIN

Query Match 39.1%; Score 604.4; DB 6; Length 6061;
Best Local Similarity 91.4%; Pred. No. 3.9e-161;
Matches 638; Conservative 0; Mismatches 60; Indels 0; Gaps 0;

QY 595 TTTACTGGCTGCTCCCTTATCGGGAAGCGGGCGCATCATATCAATGACGCGCGCTG 654
Db 3948 TTATCATGCTGGATCCAGATCTGGGCCCATCGCGCGGATCGATNNNACATGTGAG 4007
QY 655 TAAAGTGTACCTTGAGAAAGAAATTCCTGCAGCCCGCGCTTCTGGCGTTTCCATA 714
Db 4008 CAAAAGCCAGCAAAAGCCAGAACCGTAAAGAGCCGCTTCTGGCGTTTCCATA 4067
QY 715 GGCTCGCCCGCTTACGAGCATCAAAAATCGAGCTCAAGTCAGAGTGGCGAAACC 774
Db 4068 GGCTCGCCCGCTTACGAGCATCAAAAATCGAGCTCAAGTCAGAGTGGCGAAACC 4127
QY 775 CGACAGGACTATAAAGTACCAAGCGTTTCCCGCTGGAAGCTCCCTGCGCTCTCTG 834
Db 4128 CGACAGGACTATAAAGTACCAAGCGTTTCCCGCTGGAAGCTCCCTGCGCTCTCTG 4187
QY 835 TTCCGACCTTCCGCTTACCGGATACCTTCTCCCTTCTCGGGAAGCGTGGCG 894
Db 4188 TTCCGACCTTCCGCTTACCGGATACCTTCTCCCTTCTCGGGAAGCGTGGCG 4247
QY 895 TTTCTCAATGCTCAGCTGAGTATCTCAGTTCGGTGTAGTTCGCTCCAAGCTGG 954
Db 4248 TTTCTCAATGCTCAGCTGAGTATCTCAGTTCGGTGTAGTTCGCTCCAAGCTGG 4307
QY 955 GCTGTGTGACGAACCCCGTTACGCGGACGCTGCGCTTATCCGGTAACATATCGTC 1014
Db 4308 GCTGTGTGACGAACCCCGTTACGCGGACGCTGCGCTTATCCGGTAACATATCGTC 4367
QY 1015 TTGAGTCCAAACCCGGTAAGACACGACTTATCGCCACTGGCAGACCCACTGGTAACAGGA 1074
Db 4368 TTGAGTCCAAACCCGGTAAGACACGACTTATCGCCACTGGCAGACCCACTGGTAACAGGA 4427
QY 1075 TTAGCAGAGGAGTATGAGCGGTGCTACAGAGTCTTGAAGTGGTGGCCTAACTAGG 1134
Db 4428 TTAGCAGAGGAGTATGAGCGGTGCTACAGAGTCTTGAAGTGGTGGCCTAACTAGG 4487
QY 1135 GCTACACTAGAAGGACAGTATTTGGTATCTGCGCTCTGCTGAAGCCAGTTACCTTCGGA 1194
Db 4488 GCTACACTAGAAGGACAGTATTTGGTATCTGCGCTCTGCTGAAGCCAGTTACCTTCGGA 4547
QY 1195 AAAGAGTTGGTAGCTCTTGTATCCGGCAAAACACCCGCTGGTGGTGGTTTTTGG 1254
Db 4548 AAAGAGTTGGTAGCTCTTGTATCCGGCAAAACACCCGCTGGTGGTGGTTTTTGG 4607
QY 1255 TTTGCAAGCAGCAGATTACGCGCAGAAAAAAGGATCT 1292
Db 4608 TTTGCAAGCAGCAGATTACGCGCAGAAAAAAGGATCT 4645

RESULT 14
AR122286
LOCUS AR122286 6061 bp DNA PAT 16-MAY-2001
DEFINITION Sequence 6 from patent US 6165715.
ACCESSION AR122286
VERSION AR122286.1 GI:14106603
KEYWORDS

SOURCE Unknown.
ORGANISM Unknown.
REFERENCE 1 (bases 1 to 6061)
AUTHORS Collins,M.KatherineLevinge, Weiss,R.Anthony, Takeuchi,Y. and Cosset,F.
TITLE Expression systems
JOURNAL Patent: US 6165715-A 6 26-DEC-2000;
FEATURES Location/Qualifiers
source 1..6061
/organism="unknown"
BASE COUNT 1498 a 1665 c 1485 g 1405 t 8 others
ORIGIN

Query Match 39.1%; Score 604.4; DB 6; Length 6061;
Best Local Similarity 91.4%; Pred. No. 3.9e-161;
Matches 638; Conservative 0; Mismatches 60; Indels 0; Gaps 0;

QY 595 TTTACTGGCTGCTCCCTTATCGGGAAGCGGGCGCATCATATCAATGACGCGCGCTG 654
Db 3948 TTATCATGCTGGATCCAGATCTGGGCCCATCGCGCGGATCGATNNNACATGTGAG 4007
QY 655 TAAAGTGTACCTTGAGAAAGAAATTCCTGCAGCCCGCGCTTCTGGCGTTTCCATA 714
Db 4008 CAAAAGCCAGCAAAAGCCAGAACCGTAAAGAGCCGCTTCTGGCGTTTCCATA 4067
QY 715 GGCTCGCCCGCTTACGAGCATCAAAAATCGAGCTCAAGTCAGAGTGGCGAAACC 774
Db 4068 GGCTCGCCCGCTTACGAGCATCAAAAATCGAGCTCAAGTCAGAGTGGCGAAACC 4127
QY 775 CGACAGGACTATAAAGTACCAAGCGTTTCCCGCTGGAAGCTCCCTGCGCTCTCTG 834
Db 4128 CGACAGGACTATAAAGTACCAAGCGTTTCCCGCTGGAAGCTCCCTGCGCTCTCTG 4187
QY 835 TTCCGACCTTCCGCTTACCGGATACCTTCTCCCTTCTCGGGAAGCGTGGCG 894
Db 4188 TTCCGACCTTCCGCTTACCGGATACCTTCTCCCTTCTCGGGAAGCGTGGCG 4247
QY 895 TTTCTCAATGCTCAGCTGAGTATCTCAGTTCGGTGTAGTTCGCTCCAAGCTGG 954
Db 4248 TTTCTCAATGCTCAGCTGAGTATCTCAGTTCGGTGTAGTTCGCTCCAAGCTGG 4307
QY 955 GCTGTGTGACGAACCCCGCTTACGCGGACGCTGCGCTTATCCGGTAACATATCGTC 1014
Db 4308 GCTGTGTGACGAACCCCGCTTACGCGGACGCTGCGCTTATCCGGTAACATATCGTC 4367
QY 1015 TTGAGTCCAAACCCGGTAAGACACGACTTATCGCCACTGGCAGACCCACTGGTAACAGGA 1074
Db 4368 TTGAGTCCAAACCCGGTAAGACACGACTTATCGCCACTGGCAGACCCACTGGTAACAGGA 4427
QY 1075 TTAGCAGAGGAGTATGAGCGGTGCTACAGAGTCTTGAAGTGGTGGCCTAACTAGG 1134
Db 4428 TTAGCAGAGGAGTATGAGCGGTGCTACAGAGTCTTGAAGTGGTGGCCTAACTAGG 4487
QY 1135 GCTACACTAGAAGGACAGTATTTGGTATCTGCGCTCTGCTGAAGCCAGTTACCTTCGGA 1194
Db 4488 GCTACACTAGAAGGACAGTATTTGGTATCTGCGCTCTGCTGAAGCCAGTTACCTTCGGA 4547
QY 1195 AAAGAGTTGGTAGCTCTTGTATCCGGCAAAACACCCGCTGGTGGTGGTTTTTGG 1254
Db 4548 AAAGAGTTGGTAGCTCTTGTATCCGGCAAAACACCCGCTGGTGGTGGTTTTTGG 4607
QY 1255 TTTGCAAGCAGCAGATTACGCGCAGAAAAAAGGATCT 1292
Db 4608 TTTGCAAGCAGCAGATTACGCGCAGAAAAAAGGATCT 4645

RESULT 15
A60211
LOCUS A60211 6312 bp DNA circular PAT 06-MAR-1998
DEFINITION Sequence 7 from Patent WO9708330.
ACCESSION A60211

Job time: 17240 sec

```
VERSION A60211.1 GI:3715219
KEYWORDS
SOURCE
ORGANISM
REFERENCE 1 (bases 1 to 6312)
AUTHORS Collins,M.K., Weiss,R.A., Takeuchi,Y. and Cosset,F.
TITLE EXPRESSION SYSTEMS
JOURNAL Patent: WO 9708330-A 7 06-MAR-1997;
CANCER RES CAMPAIGN TECH (GB)
FEATURES
source
location/Qualifiers
1..6312
/organism="unidentified"
/db_xref="taxon:32644"
BASE COUNT 1575 a 1692 c 1561 g 1476 t 8 others
ORIGIN

Query Match 39.1%; Score 604.4; DB 6; Length 6312;
Best Local Similarity 91.4%; Pred. No. 3.9e-161;
Matches 638; Conservative 0; Mismatches 60; Indels 0; Gaps 0;

QY 595 TTTACTGGCTGCTCCCTTATCGGGAGCGGGGGCGATCATATCAATGACGCCCGCTG 654
DB 4199 TTATCATGTCTGGATCCAGATCTGGGCCATCGCGCGCGATCGATNNNNACATGTGAG 4258
QY 655 TAAAGTGTACGTGAGAAAGAAATTCCTCGAGCCCGCGCTGCTGCGCTTTTCCATA 714
DB 4259 CAAAGCCAGCAAAAGCCAGGACCGTAAAGAGCGCGGTGCTGCGCTTTTCCATA 4318
QY 715 GGCTCCGCCCCCTGACGAGCATCACAAAATCGACGCTCAAAGTCAGAGGTGGCGAAACC 774
DB 4319 GGCTCCGCCCCCTGACGAGCATCACAAAATCGACGCTCAAAGTCAGAGGTGGCGAAACC 4378
QY 775 CGACAGACTATAAAGATACAGCGCTTCCCGCTGGAAGCTCCCTGCTGCGCTCTCCTG 834
DB 4379 CGACAGACTATAAAGATACAGCGCTTCCCGCTGGAAGCTCCCTGCTGCGCTCTCCTG 4438
QY 835 TTCGAGCCCTGCGCTTACCGGATACCTGTCGCGCTTTCCTCCCTTCGCGGAGCGTGCGC 894
DB 4439 TTCGAGCCCTGCGCTTACCGGATACCTGTCGCGCTTTCCTCCCTTCGCGGAGCGTGCGC 4498
QY 895 TTTCTCAATGCTCAGCGTGTAGGTATCTCAGTTGCGGTGTAGTTCGCTCCAAAGCTGG 954
DB 4499 TTTCTCAATGCTCAGCGTGTAGGTATCTCAGTTGCGGTGTAGTTCGCTCCAAAGCTGG 4558
QY 955 GCTGTGTGACGACACCCCGCTTACGCGCCGACCGCTCGCGCTTATCCGGTAACCTATCGTC 1014
DB 4559 GCTGTGTGACGACACCCCGCTTACGCGCCGACCGCTCGCGCTTATCCGGTAACCTATCGTC 4618
QY 1015 TTGAGTCCAAACCCGGTAAGACAGACTTATCGCCACTGGCAGCAGCCACTGGTAACAGGA 1074
DB 4619 TTGAGTCCAAACCCGGTAAGACAGACTTATCGCCACTGGCAGCAGCCACTGGTAACAGGA 4678
QY 1075 TTAGCAGAGCGAGGTATGTAGGCGGTGCTACAGAGTTCTTGAAGTGTGGCTTAACACG 1134
DB 4679 TTAGCAGAGCGAGGTATGTAGGCGGTGCTACAGAGTTCTTGAAGTGTGGCTTAACACG 4738
QY 1135 GCTACACTAGAAGACAGTATTTGGTATCTGCGCTCTGCTGAAGCCAGTACCTTCGGAA 1194
DB 4739 GCTACACTAGAAGACAGTATTTGGTATCTGCGCTCTGCTGAAGCCAGTACCTTCGGAA 4798
QY 1195 AAAGAGTTGTAGCTCTTGTATCCGGCAACAAACCACCGCTGTAGCGGTGTTTTTTG 1254
DB 4799 AAAGAGTTGTAGCTCTTGTATCCGGCAACAAACCACCGCTGTAGCGGTGTTTTTTG 4858
QY 1255 TTTGCAAGCAGCAGATTACGCGCAGAAAAAAGGATCT 1292
DB 4859 TTTGCAAGCAGCAGATTACGCGCAGAAAAAAGGATCT 4896
```

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: January 17, 2002, 12:03:54 ; Search time 599.86 Seconds
(without alignments)
2210.986 Million cell updates/sec

Title: US-09-242-202A-27
Perfect score: 1547
Sequence: 1 GGTACCTGCCACCATGGCGC.....CTGCACAGCGCTCTCCACA 1547

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 930621 seqs, 428662619 residues

Total number of hits satisfying chosen parameters: 1861242

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : N_Geneseq_1101.*
1: /SIDS2/gcgdata/geneseq/geneseq/NA1980.DAT.*
2: /SIDS2/gcgdata/geneseq/geneseq/NA1981.DAT.*
3: /SIDS2/gcgdata/geneseq/geneseq/NA1982.DAT.*
4: /SIDS2/gcgdata/geneseq/geneseq/NA1983.DAT.*
5: /SIDS2/gcgdata/geneseq/geneseq/NA1984.DAT.*
6: /SIDS2/gcgdata/geneseq/geneseq/NA1985.DAT.*
7: /SIDS2/gcgdata/geneseq/geneseq/NA1986.DAT.*
8: /SIDS2/gcgdata/geneseq/geneseq/NA1987.DAT.*
9: /SIDS2/gcgdata/geneseq/geneseq/NA1988.DAT.*
10: /SIDS2/gcgdata/geneseq/geneseq/NA1989.DAT.*
11: /SIDS2/gcgdata/geneseq/geneseq/NA1990.DAT.*
12: /SIDS2/gcgdata/geneseq/geneseq/NA1991.DAT.*
13: /SIDS2/gcgdata/geneseq/geneseq/NA1992.DAT.*
14: /SIDS2/gcgdata/geneseq/geneseq/NA1993.DAT.*
15: /SIDS2/gcgdata/geneseq/geneseq/NA1994.DAT.*
16: /SIDS2/gcgdata/geneseq/geneseq/NA1995.DAT.*
17: /SIDS2/gcgdata/geneseq/geneseq/NA1996.DAT.*
18: /SIDS2/gcgdata/geneseq/geneseq/NA1997.DAT.*
19: /SIDS2/gcgdata/geneseq/geneseq/NA1998.DAT.*
20: /SIDS2/gcgdata/geneseq/geneseq/NA1999.DAT.*
21: /SIDS2/gcgdata/geneseq/geneseq/NA2000.DAT.*
22: /SIDS2/gcgdata/geneseq/geneseq/NA2001.DAT.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1547	100.0	1547	19	Humanised vector p
2	1310.4	84.7	1807	19	Humanised vector p
3	1086.6	70.2	2308	19	Humanised vector p
C 4	755	48.8	2932	13	COS cell expressio
C 5	755	48.8	2932	17	Expression vector
C 6	755	48.8	2932	17	Vector plasmid piH
C 7	755	48.8	2932	19	Expression vector
C 8	755	48.8	2932	20	Expression vector
C 9	755	48.8	2932	20	Eukaryotic cell ex
C 10	755	48.8	2932	22	Expression vector
11	607.4	39.3	7252	20	Complete sequence

12	604.4	39.1	5865	18	AAT90692	Plasmid FBdelPRDSA
13	604.4	39.1	6028	18	AAT90689	Plasmid FBdelPASAF
14	604.4	39.1	6061	18	AAT90690	Plasmid FBdelPMOSA
15	604.4	39.1	6312	18	AAT90691	Plasmid FBdelPGASA
16	604.2	39.1	7073	18	AAT50962	TF8-569 CDR-grafte
17	604.2	39.1	7864	18	AAT50963	TF8-569 CDR-grafte
18	604.2	39.1	11795	21	AAAS9345	Nucleotide sequenc
19	604.2	39.1	13254	17	AAT40915	Nucleotide sequenc
20	604.2	39.1	13254	21	AAAS2151	pEel2 Combo BM 12
C 21	604.2	39.1	13254	21	AAAS2151	Complete nucleotid
22	603	39.0	608	19	AAV21731	ColE1 origin of re
23	603	39.0	2077	20	AAZ23771	Vector pASlib DNA.
24	603	39.0	2462	21	AAAT74638	Plasmid pSP72. Es
25	603	39.0	2577	14	AAQ36620	Expression plasmid
26	603	39.0	2704	20	AAZ06758	Vector pUC28 nucle
C 27	603	39.0	2927	17	AAT29158	Plasmid pTRP. Syn
C 28	603	39.0	2927	17	AAT31789	Plasmid pTRP. Syn
C 29	603	39.0	2939	16	AAQ87350	Plasmid BGINV. Sy
C 30	603	39.0	2939	19	AAV37293	pBGINV plasmid use
31	603	39.0	3003	11	AAQ05745	Plasmid P.L-mu-smc
32	603	39.0	3104	10	AAAN90296	DNA target sequenc
33	603	39.0	3130	19	AAV46334	Universal reporter
C 34	603	39.0	3175	20	AAZ20066	DNA encoding gluta
35	603	39.0	3327	22	AAAC86263	Plasmid GHRH1-29YA
C 36	603	39.0	3331	21	AAQ84434	pBC66-01, containi
C 37	603	39.0	3331	21	AAZ99245	Nucleotide sequenc
C 38	603	39.0	3344	21	AAAS9495	Transgenic unc-119
C 39	603	39.0	3369	22	AAAC86255	pGHRH-4 4A5k const
40	603	39.0	3403	9	AAAN80956	Plasmid pDS5/RBSII
41	603	39.0	3414	12	AAQ12785	pDS56/RBSII-2 sequ
42	603	39.0	3414	14	AAQ34609	Expression plasmid
43	603	39.0	3415	12	AAQ12784	pDS56/RBSII-1 sequ
44	603	39.0	3415	14	AAQ34608	Expression plasmid
45	603	39.0	3416	9	AAAN80958	Plasmid pDS56/RBSI

ALIGNMENTS

RESULT 1	
AAV21732	
ID	AAV21732 standard; cDNA; 1547 BP.
XX	
AC	AAV21732;
XX	
DT	17-AUG-1998 (first entry)
XX	
DE	Humanised vector pITL-A.
XX	
KW	Vector; vaccine; tumour; antigen; plasmid pITL-A; ds.
OS	Chimeric - Homo sapiens.
OS	Chimeric - Escherichia coli.
XX	Synthetic.
XX	
FT	Key
FT	CDS
FT	Location/Qualifiers
FT	472..680
FT	/*tag= c
FT	/note= "SupF gene"
FT	686..1292
FT	/*tag= d
FT	/note= "ColE1 origin of replication"
FT	1311..1547
FT	/*tag= e
FT	/note= "RANTES promoter"
XX	
PN	WO9806863-A1.
XX	
PD	19-FEB-1998.
XX	
PF	14-AUG-1997; 97WO-US14306.
XX	
PR	14-AUG-1996; 96US-0023931.

XX (USSH) US DEPT HEALTH & HUMAN SERVICES.
PA Nelson EL, Nelson PJ;
PI WPI; 1998-159552/14.
DR Humanised polynucleotide vectors - comprising human derived promoter
XX and sequence acceptance site, used for the production of vaccines
PT Claim 14; Page 56-57; 125pp; English.
PS Plasmid pITL-A comprises a base vector for novel humanised
XX polynucleotide vectors. The vector in Escherichia coli DH
CC 10-beta/P3 is deposited as ATCC 98401. Novel vectors comprise a
CC human-derived promoter or mammalian homologue which is functional
CC in mammalian target tissue and cells and a sequence acceptance site
CC (see AAV21735-36) which accepts cDNA products from RT-PCR cloning.
CC They also contain minimal non-human components, such as a replication
CC origin (see AAV21715) and selectable marker gene (see AAV21717-18) that
CC are necessary for production of the vector. The novel vectors are
CC used to express target antigens, especially tumour antigens. They
CC are non-replicating in mammalian cells but are capable of extended
CC stable expression of target sequences generating an immune response
CC in immunised individuals. The vectors selectively elicit immune
CC responses to the target sequences with little or no immune response
CC to the other components of the vectors. The target antigens are
CC expressed as intracellular polypeptides or peptides and, as such,
CC are processed as self polypeptides or peptides and appropriately
CC presented on antigen presenting cells.
XX SQ Sequence 1547 BP; 344 A; 408 C; 425 G; 370 T; 0 other;

Query Match 100.0%; Score 1547; DB 19; Length 1547;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1547; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGTACCTGCCACCATGGCGGATCTTTATCACTGATAAGTTGGTGACATATTATGTT 60
DB 1 ggtacctgccaccatgscggatctttatcactgataaagttggtggacatatattg 60

QY 61 TATCACTGATTAAGTCTCAAGCATGACAAAGTTGACGCGAATACAGTATCGTGCCGG 120
DB 61 tatcactgataaagtctcaagcatgacaaagtgtgacgcgaatacacgtgacgtgccg 120

QY 121 CCTGTGACTGTTCAACGAGGTGCGGTAGACGCTGACGACACGCAAACTGSCGGAACG 180
DB 121 cctgtgactgtttcaacgaggtgcggttagacgctgacgacacgcaaaactgscggaacg 180

QY 181 GTTGGGGGTGACGACGCGCGCTTTTACTGGCACTTTCAGGAACAAAGCGCGCTTTAAGG 240
DB 181 gttgggggtgacgacgscgcttttactggcacttcaggaacaaagcgcgcttaagg 240

QY 241 GCCATATGTTGATGATGCTTACCCGAGCGGGGATGGGGGAGACCTGTAGTCAGAG 300
DB 241 gccatatggtgagtgtgcttgacccagcgcggtgagggagacacctgtagtccagag 300

QY 301 CCCCGGGGACGACAGCCCAATGCGCTCTTCCCTGCGAGGATGAGTAGTGAGTCGCTC 360
DB 301 ccccggggacgacagcccaatgscgctcttccctgcgaggtgagtagtgagtcgctc 360

QY 361 TCTGTCCTTGAAGTTGCACTTCCAGTCCACCGCTTGTCTTAATAAATAAATTAAGTT 420
DB 361 tctgctccttgaagttgcaactccagtcacagcctgttcctaataaataaattagtt 420

QY 421 GCATCATTTTGTCTGACTAGGTGCTCTATATATATATATATATATATATATATATAT 480
DB 421 gcatcatitttgtctgactaggtgtctctatataataataataataataataataattctt 480

QY 481 CGGACTTTTGAAGTATGTTGTTGGGGGAGGATTCGACCTTCGAAGTCGATGACGGC 540
DB 481 cggacttttgaagtgtgtgtgggggagggatctcgacaccttcgaagtcgatgacggc 540

RESULT 2
AAV21733
ID AAV21733 standard; cDNA; 1807 BP.

QY 541 AGATTAGAGTCTGCTCCCTTTGGCCGCTCGGAACCCACACCGGTAATGCTTTTACT 600
DB 541 agatttagagctgctccctttggccgctcggaaacccacacccggtaagtctttact 600

QY 601 GGCCTGCTCCCTTATTCGGAAGCGGGCGCATCATATCAATGACGCGCGCTGTAAAGT 660
DB 601 ggcctgctcccttatttcggaagcgggcgcatcatacaatgacgcgcgctgtaaagt 660

QY 661 GTTACGTTGAGAAAGAAATTCCTGCAGCGCGCGCTTGTGCGCTTTTTCATAGGCTCC 720
DB 661 gttacgttgagaaagaattcctgcagcgcgcgtgtgctgcgctttttccatagctcc 720

QY 721 GCCCCCTTACGAGCATCACAAAATCGACGCTCAAGTCAGAGTGGGCAACCCGACAG 780
DB 721 gcccccctgaagcagcatcacaaaaatcgacgctcaagtgcagaggtggcgaaacccgacag 780

QY 781 GACTATAAAGATACGAGGCTTCCCTCGGAAGCTCCCTCGTGGCTCTCCTGTTCCGA 840
DB 781 gactataaagatacaggcgctttcccccctggaaagctccctcgtgctcctcctgttccga 840

QY 841 CCTGCGCGCTTACCGGATACCTGTCCGCTTTCTCCCTTCGGGAAGCGTGGCGCTTTC 900
DB 841 cctgcgcgcttaccggatacctgtccgctttctcccttcgggaagcgtggcgcttctc 900

QY 901 AATGCTACGCTGTAGGTATCTCAGTTGGGTAGGTGCTTGGTCCCAAGCTTGGGCTGTG 960
DB 901 aatgctcacgctgtaggtatctcagttcgtgttaggtcgttcctccaagctggcgctgtg 960

QY 961 TGCAGCAACCCCGCTTACGCGCGCTGCGCTTATTCGCGTAACTATCGTCTTTGAGT 1020
DB 961 tgcagcaaccccgcttaccggcgctgacgcgcgtgctccttcccgtaactatcgtcttgagt 1020

QY 1021 CCAACCCGCTAAGACACGACTTATCGCCACTGGCAGACGCCACTGGTAACAGGATTAGCA 1080
DB 1021 ccaacccgctaagacacgacttattcgcaactggcagcagcactggtaaacaggattagca 1080

QY 1081 GAGCAGAGTATGAGCGGTGTACAGAGTTCTTGAAGTGGTGGCTTAAGTACGCGCTACA 1140
DB 1081 gagcagagtatgagcggtgtacagagttcttgaagtgggtggcctaactacggtaca 1140

QY 1141 CTAGAGGACAGTATTTGTTGTTATCTGGCTCTGCTGAAGCCAGTTACTTTCGGAAGAGAG 1200
DB 1141 ctagaggaagcagttatttggtaattctgctcgtcgtgaagcaggttaaccttcggaaaaagag 1200

QY 1201 TTGGTAGCTTTGATCCGCCAAACAAACACCCGCTGGTAGCGGTGGTGTGTTTGTGGA 1260
DB 1201 ttggtagctttgatccggcgaacaaacccacccgctggtgagcgtgtttttttgttgca 1260

QY 1261 AGCAGCAGATTAGCGCGCAAAAAGGATCTGGGGGATCCGGGAGAGCTCACTCTAGATG 1320
DB 1261 agcagcagatttagcgcgcaaaaaaggatctgggggatccggagagactcactctagatg 1320

QY 1321 AGAGAGCAGTGGGGAGACAGAGACTCGAATTTCCGAGAGCTATTTTTCAGTTTCTTTTTC 1380
DB 1321 agagagcagtggggagagacagagactcgaatttccggagcattattcagttttcttcc 1380

QY 1381 CGTTTTGTGCAATTTTCACTTATGATACCGCCCAATGCTTGGTGTGCTTATTTGGAAGTCC 1440
DB 1381 cgttttgtgcaattttcaacttatgataccggccaatgctgtgttctatttttgaaactcc 1440

QY 1441 CCTTAGGGGATGCCCTCAACTGGGCCCTATTAAGGCCAGCTGAGCTGCAGAGGATTCC 1500
DB 1441 ccttaggggatgccctcaactggccctataaaggccagcctgagctgagctgagaggattcc 1500

QY 1501 TGCAGAGGATCAAGACAGCAGCTGGACCTCGCACACGCTCTCCACCA 1547
DB 1501 tgcagaggatcaagacagcagcgtggacctcgacacgacctctcccca 1547

XX AAV21733;
 AC
 XX
 DT 17-AUG-1998 (first entry)
 XX
 DE Humanised vector pITL-1.
 XX
 KW Vector; vaccine; tumour; antigen; plasmid pITL-1; ds.
 XX
 OS Chimeric - Homo sapiens.
 OS Chimeric - Escherichia coli.
 OS Synthetic.
 XX
 XX Location/Qualifiers
 FH Key
 FT CDS
 FT
 FT misc_feature
 FT
 FT promoter
 FT
 FT
 FT
 FT
 XX W09806863-A1.
 PN
 XX
 XX 19-FEB-1998.
 XX
 PF 14-AUG-1997; 97WO-US14306.
 XX
 PR 14-AUG-1996; 96US-0023931.
 XX
 PA (USSH) US DEPT HEALTH & HUMAN SERVICES.
 XX
 PI Nelson EL, Nelson PJ;
 XX
 DR WPI; 1998-159552/14.
 XX
 XX Humanised polynucleotide vectors - comprising human derived promoter
 PT and sequence acceptance site, used for the production of vaccines
 XX
 PS Claim 14; Page 56; 125pp; English.
 CC
 CC Plasmid pITL-1 comprises a base vector for novel humanised
 CC polynucleotide vectors. The vector in Escherichia coli DH
 CC 10-beta/P3 is deposited as ATCC 98400. Novel vectors comprise a
 CC human-derived promoter or mammalian homologue which is functional
 CC in mammalian target tissue and cells and a sequence acceptance site
 CC (see AAV21735-36) which accepts cDNA products from RT-PCR cloning.
 CC They also contain minimal non-human components, such as a replication
 CC origin (see AAV21715) and selectable marker gene (see AAV21717-18) that
 CC are necessary for production of the vector. The novel vectors are
 CC used to express target antigens, especially tumour antigens. They
 CC are non-replicating in mammalian cells but are capable of extended
 CC stable expression of target sequences generating an immune response
 CC in immunised individuals. The vectors selectively elicit immune
 CC responses to the target sequences with little or no immune response
 CC to the other components of the vectors. The target antigens are
 CC expressed as intracellular polypeptides or peptides and, as such,
 CC are processed as self polypeptides or peptides and appropriately
 CC presented on antigen presenting cells.
 XX
 SQ Sequence 1807 BP; 419 A; 463 C; 492 G; 433 T; 0 other;

Query Match 84.7%; Score 1310.4; DB 19; Length 1807;
 Best Local Similarity 99.9%; Pred. No. 0;
 Matches 1311; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 GGTACCTGCCACCATGGCGGGATCTCTTATCACTGATGAAGTGGTGGACATATATGTT 60
 DB
 1 ggtacctgccaccatggcgcggattcttctactcgtatgaagtgggtggacattattgtt 60

QY 61 TATCAGTGATAAAGTGTCAAGCATGTGCAAAAGTTGACGCCGAATACAGTGCCTGCGCG 120
 DB
 61 tatcaagtataaagtgtaagcatgacaaagtgcagccgaatacagtgatcgtgcg 120
 QY 121 CCCTGGACTCTTGAACGAGGTGCGGTGAGCGTCTGACGACACGCAAACTGGCGGAACG 180
 DB
 121 ccttggactgttgaacgaggtcgcgtagacggtctgacgacacgcgaactggcggaacg 180
 QY 181 GTTGGGGTGCACGACGCCGGCTTTACTGGCACTTTCAGGAACAGCGGCGCTTAAGG 240
 DB
 181 gttgggggtgcagcagccggcgtttactggcacttccaggaacagcggcgcttaagg 240
 QY 241 GCCATATGTTGAGTGGATGCTTACCCAGCGGGGATGGGGAGACCTCTAGTACAGAG 300
 DB
 241 gccatatgtgagtggatgcttgaccccgagcgggatggggagacctgtagtcaag 300
 QY 301 CCCCCGGGACGACAGGCCCAATGCCCTCTTCCCTTGCAGGATGAGTAGTGCGCTC 360
 DB
 301 cccccgggcagcacaggccaatgccgtccttcccttcaggtagtagtgagtgccctc 360
 QY 361 TCCTGGCCCTGGAGTTGCCACTCCAGTCCGACGACGCTTGCTCTAATAAATAAAGTT 420
 DB
 361 tcctggccctgggaagttgcccactccagtgccaccagccttgcctataaaattaa 420
 QY 421 GCATCATTTTGTCTGACTAGTGTCTCTATAATATTATTAAGCTTGATATCGAATTCTTT 480
 DB
 421 gcatcatttgtctgactaggtgtccctctataataattataagcttgatatacgaattct 480
 QY 481 CGGACTTTTGAAGTGATGTTGGGGAGGATTCGAACCTTCGAAGTCGATGACGGC 540
 DB
 481 cggacttttgaagtgatggtggggagaggttcgaaacctcgaagtcgagtcgagcgc 540
 QY 541 AGATTTAGAGTCTGCTCCCTTTGGCGCTCGGGAACCCACCGGTAATGCTTTACT 600
 DB
 541 agatttagagtcgtccctttggcgcctcgggaaacccaccacgggtaattgctttact 600
 QY 601 GGCTGTCTCTTATCGGGAAGCGGCGCATCATATCAATGACGCGCGCTGTAAAGT 660
 DB
 601 ggctgtctcttattcggaagcgggcgcatcatacaaatgacgcgcgctgtaaagt 660
 QY 661 GTTACGTTGAGAAAGATTCTCGACGCCCGCGGTGTGCGGCTTTTCCATAGGCTCC 720
 DB
 661 gttacgttgagaaagaattcctgcagccgcgcgtgtgctggttttccataggtctcc 720
 QY 721 GCCCCCTGACGAGCATCACAAAATCGACGCTCAAGTCAGAGTGGGGAACCCGACAG 780
 DB
 721 gcccccctgacgagcatcaaaaaatcgacgctcaagtcaaggtgagcgaacccgcag 780
 QY 781 GACTATAAAGATACAGCGCTTTCCCTCTGGAAGCTCCCTCGTGGCTCTCTCTGTCGA 840
 DB
 781 gactataaagatacaccagcgcttcccccctggaagctccctgctgctctctctccga 840
 QY 841 CCCGTCGCTTACCGGATACCTGTGCGCTTTTCTCCCTTGGGAAGCGTGGCGCTTCTC 900
 DB
 841 cccgtcgcgttacccgataacctgccgcttcccttccttcctcggaagcgtggcgcttctc 900
 QY 901 AATGCTCACGCTGTAGGTATCTCAGTTGGGTGAGTGGTTCGCTCCCAAGTGGGCTGTG 960
 DB
 901 aatgctcacgctgtaggtatctctcagttcgggtgtaggttcgctcctcaagctggcgctg 960
 QY 961 TGACGAACCCCGCTTTCAGCGCCGACGCTGCGCTTTATCCGGTAACTATATGCTGTGAGT 1020
 DB
 961 tgaacgaaccccccggttcagcccgacgcgtgcgcttatcccgtaactatcgtttgagt 1020
 QY 1021 CCAACCCGGTAAAGACAGACTTATCGCCACTTGGCAGCAGCCACCTGGTAAACAGGATTAGCA 1080
 DB
 1021 ccaacccggtaagacacagacttatcgcaactggcagcagcactggtaaacaggattagca 1080
 QY 1081 GAGCGAGGTATGAGGCGGTGCTACAGATTCCTTGAAGTGGTGGCGCTTAACCTACCGCTACA 1140
 DB
 1081 gagcgaggtatgagcggtgctacagagttcttgaagtgggtccctaactacggctaca 1140
 QY 1141 CTAGAAGGACAGTATTGTTGGTATCTGCGCTCTGCTGAAGCCAGTTACCTTCGGAAGAAAGAG 1200

||||| ctagaagacagtagtattgtgtatcgcgtctgctgaagccagttacccttcggaaaaagag 1200
Db 1141
QY 1201 TTGGTAGCTCTTGATCGGCAAAACACCGCTGGTAGCGGTGGTTTGTGTTGCA 1260
Db 1201 ttggtagctcttgatccggcaacacacccgctggtagcgggtgtttttttttttgttga 1260
QY 1261 AGCAGCAGATTACGCGCAAAAGAGATCTGGGGATCGCGAGAGCTCAC 1312
Db 1261 agcagcagattacgcagaaaaaagatctggggatccggagagctccc 1312
RESULT 3
AAV21734
ID AAV21734 standard; cDNA: 2308 BP.
XX AC AAV21734;
XX AC AAV21734;
DT 17-AUG-1998 (first entry)
XX Humanised vector pITL-1 GFP.
XX Vector; vaccine; tumour; antigen; plasmid pITL-1 GFP;
KW green fluorescent protein; ds.
XX Chimeric - Homo sapiens.
OS Chimeric - Escherichia coli.
OS Chimeric - Aequorea victoria.
OS Synthetic.
XX Key Location/Qualifiers
FH 20..734
FT CDS
FT /*tag= a
FT /product= green fluorescent protein
FT 973..1181
FT /*tag= c
FT /note= "SupF gene"
FT 1191..1793
FT /*tag= d
FT /note= "ColE1 origin of replication"
FT 2063..2308
FT /*tag= e
FT /note= "RANTES promoter"
XX WO9806863-A1.
XX 19-FEB-1998.
XX 14-AUG-1997; 97WO-US14306.
XX 14-AUG-1996; 96US-0023931.
XX (USSH) US DEPT HEALTH & HUMAN SERVICES.
XX Nelson EL, Nelson PJ;
XX WPI; 1998-159552/14.
XX Humanised polynucleotide vectors - comprising human derived promoter
PT and sequence acceptance site, used for the production of vaccines
XX Example 11; Page 57-58; 125pp; English.
XX Plasmid pITL-1 GFP comprises base vector pITL-1 (see AAV21733) and
CC a humanised green fluorescent protein (GFP) reporter sequence
CC (see AAV21725). Novel humanised vectors comprise a human-derived
CC promoter or mammalian homologue which is functional in mammalian
CC target tissue and cells and a sequence acceptance site which
CC accepts cDNA products from RT-PCR cloning. The novel vectors are
CC used to express target antigens, especially tumour antigens. They
CC are non-replicating in mammalian cells but are capable of extended
CC stable expression of target sequences generating an immune response
CC in immunised individuals. The vectors selectively elicit immune

CC responses to the target sequences with little or no immune response
CC to the other components of the vectors. The target antigens are
CC expressed as intracellular polypeptides or peptides and, as such,
CC are processed as self polypeptides or peptides and appropriately
CC presented on antigen presenting cells.
XX Sequence 2308 BP; 571 A; 601 C; 604 G; 532 T; 0 other;
SQ
Query Match 70.2%; Score 1086.6; DB 19; Length 2308;
Best Local Similarity 99.6%; Pred. No. 0;
Matches 1089; Conservative 0; Mismatches 4; Indels 0; Gaps 0;
QY 220 GAACAAGCGCGGCTTAAGGGCCATATGGTCAGTGGATGCTTACCCAGCGGGAT 279
Db 721 gtacaagtgcgccttaaggccatattgtgagtgagtccttgaccccaaggcgggat 780
QY 280 GGGGAGACCTGTAGTCAGAGCCCGGGCAGCAGACAGGCCAATGCCCTTCCCTG 339
Db 781 gggggagacctgtagcagagcccccggcagacagccaatgccgtccctccctgc 840
QY 340 AGGATGAGTGTAGTGCCTCTCTGGCCCTGGAAGTTGCCACTCCAGTGCCACAGCC 399
Db 841 aggatgagttagtgctctcctggccctggaagtggccactccagtggccaccagcc 900
QY 400 TTGTCTTAATAAATTAAAGTTGCATCATTTTCTGACTAGGTGCTCTATAATATTAT 459
Db 901 ttgtcctaataaattaaagttagcatatttctgactagtgctcctataattattat 960
QY 460 AAGCTTGATATCGAATTCTTTCGGACTTTTGAAGTGTAGTGGTGGGGAAGGATTCGA 519
Db 961 aagcttgatatacgaattcttcgacttttgaagtgtggtggggaagattcga 1020
QY 520 ACCTTCGAAGTCGATGACGGCAGATTTAGAGTCTGCTCCCTTTGGCCGCTCGGAAACCCC 579
Db 1021 accttcgaagtcgatgcagggcagatttagagtcgtcctcctttggcgtcggaaaccc 1080
QY 580 ACCACGGGTAAAGCTTTTACTGGCTGCTCCCTTATCGGGAAGCGGGGCATCATATCA 639
Db 1081 accacgggtaagtctttactggcctgtcctcttatcgggaagcggggcgatcatatca 1140
QY 640 AATGACGGCGCGCTGTAAAGTGTAGTTGAGAAGAATTCCTGCAGCCCGCGCTTC 699
Db 1141 aatgacgcgcgcgtgtaaagtgttaagttagaagaattcctgcagccgcgcgcttgc 1200
QY 700 TGGCGTTTTCATAGGCTCCGCCCTCCCTGAGAGCATCACAAAATTCAGGCTCAAGTC 759
Db 1201 tggcgttttccataggtccgcggccctgcagcagcatcacaaaaatcgacgtcaagtc 1260
QY 760 AGAGGTGGCGAAACCCGACAGGACTATAAGATACCAGCGCTTCCCTCGGAAGCTCC 819
Db 1261 agaggtggcgaaacccgcagagactataaagataccaggcgtttccctctggagctccc 1320
QY 820 TCGTGGCTCTCTGTTCCGACCGCTGCCCTTACCGGATACCTGTCCGCTTCTCCCTT 879
Db 1321 tcgtgcgtctcctgcttcgcagccctgcgcttaccgataacctgtccgctttccctt 1380
QY 880 CGGGAAGCTGGCGCTTCTCAATGCTCAGCTGAGGTATCTCAGTTGGGCTAGGTGCG 939
Db 1381 cgggaagcgtggcgctttctcaatgctcagcgtgtaggtatctcagttcgggtgtagtgc 1440
QY 940 TTCGCTCAAGCTGGGCTGTGTCAGGAACCCCGCTTCAGCCCGACCGCTGCGCTTAT 999
Db 1441 tcgctccaagctgggctgtgtgcaggaacccccccttcagcccccgcgcgtcgttat 1500
QY 1000 CGGTAACCTATGCTGTGAGTCCAAACCCGTAAGACACACTTATGCCACTGGCAGCAG 1059
Db 1501 cggtaactatgcttctgagtcacaaccccggaagacacgacttatgcactggcagcag 1560
QY 1060 CCACTGGTAACAGGATTACGAGCGAGGTATGTAGCGGTGCTACAGAGTTCTTGAAGT 1119
Db 1561 ccactggtaacaggattagcagcaggtatgttagcgggtgctacagagattcttgaagt 1620

QY 1120 GGTGGCTTAACGCTACACTAGAAGACACAGTATTGGTATCTCGCTCTGCTGAAGC 1179
 Db 1621 ggtggcctaactacgctacactagaagacagatttggatctcgtctgctgaagc 1680
 QY 1180 CAGTTACCTTCGGAAGAGTTGGTAGCTCTTGATCCGGCAAAACACCGCTGGTA 1239
 Db 1681 cagttaccttcggaagagttgtagctcttgatccggcaaaacacacccgctggta 1740
 QY 1240 GGGGTGGTTTTTGTTCGACGACGACATTACGCGCAGAAAAAAGGATCTGGGGGAT 1299
 Db 1741 gcgggtggTTTTTGTTCGACGACGACATTACGCGCAGAAAAAAGGATCTGGGGGAT 1800
 QY 1300 CCGGAGAGCTCAC 1312
 Db 1801 ccggagagctccc 1813
 RESULT 4
 ID AAQ21163/c
 XX AAQ21163;
 XX 21-MAY-1992 (first entry)
 XX COS cell expression vector pIH3.
 XX cloning technique; CD2; cell surface antigen; immunodiagnosis;
 KW tumour; ss.
 XX Synthetic.
 FH Key Location/Qualifiers
 misc_feature 1..589 /tag= a
 /function= ori
 /note= "derived from pMB1 origin"
 590..597 /tag= b
 /function= linker
 /note= "SacII linker"
 598..799 /tag= c
 /label= supF_gene
 /note= "derived from synthetic tyrosine suppressor
 tRNA gene"
 800..947 /tag= d
 /label= 948..1500 /note= "remnant of ASV LTR fragment (PvuII to MluI)"
 /tag= e
 /function= enhancer
 /note= "derived from cytomegalovirus AD169 enhancer"
 1501..1650 /tag= f
 /note= "derived from HIV TATA and tat responsive-
 elements"
 1651..1716 /tag= g
 /function= polylinker
 /note= "derived from pILXAN polylinker"
 2570..2917 /tag= h
 /function= ori
 /note= "derived from SV40 origin of replication"
 2918..2922 /tag= i
 /note= "derived from pIVX, remnant of RI site
 from polylinker"
 1717..2569 /tag= j
 /note= "derived from psv to splice and

FT poly-A addition signals*
 XX
 PN W09201049-A.
 XX
 PD 23-JAN-1992.
 XX
 PF 15-JUL-1990; 90WO-US04986.
 XX
 PR 13-JUL-1990; 90US-0553759.
 XX
 PA (GEO-) GEN HOSPITAL CORP.
 XX
 PI Seed B, Aruffo A, Amiot M;
 XX WPI; 1992-056864/07.
 DR
 XX
 PT New CD53 cell surface antigen and DNA encoding it - for
 PT immuno-therapy and diagnosis of haematopoietic neoplasms, etc.
 XX
 PS Disclosure; Fig 1; 160pp; English.
 XX
 CC This COS cell expression vector was constructed from pISV
 CC (Little et al., Mol. Biol. Med. 1:473-488 (1983)) by inserting a
 CC synthetic transcription unit between the suppressor tRNA gene and
 CC the SV40 origin. The transcription unit consisted of a chimeric
 CC promoter composed of human cytomegalovirus AD169 immediate early
 CC enhancer sequences fused to the HIV LTR -67 to (plus)80 sequences.
 CC A polylinker contg. 2 BstXI sites separated by a 350bp stuffer was
 CC inserted immediately downstream from the LTR (plus)80 sequence.
 CC Downstream from the polylinker were placed the SV40 small
 CC t antigen splice and early region polyadenylation signals derived
 CC from psV2. The new vector, which forms part of the wider disclosure
 CC of the specification, was used in the cloning of the human CD2
 CC antigen gene.
 XX
 SQ Sequence 2932 BP; 775 A; 698 C; 682 G; 777 T; 0 other;
 Query Match 48.8%; Score 755; DB 13; Length 2932;
 Best Local Similarity 97.6%; Pred. No. 3.9e-219;
 Matches 788; Conservative 0; Mismatches 15; Indels 4; Gaps 2;
 QY 469 ATCGAATTCCTTCGAGACTTTTGAAGTGTGGTGGGGAAGGATTCGAACCTTCGAA 528
 Db 804 AGCAGATTCTTCGACTTTTGAAGTGTGGTGGGGAAGGATTCGAACCTTCGAA 745
 QY 529 GTCGATGACGGCAGATTTAGAGTCTGCTCCCTTTGGCCGCTCGGGAAACCCACACGGGT 588
 Db 744 GTCGATGACGGCAGATTTAGAGTCTGCTCCCTTTGGCCGCTCGGGAAACCCACACGGGT 685
 QY 589 AATGCTTTTACTGCGCTGCTCCCTTATCGGGAAGCGGGCGCATCATCAATGACGCG 648
 Db 684 AATGCTTTTACTGCGCTGCTCCCTTATCGGGAAGCGGGCGCATCATCAATGACGCG 625
 QY 649 CCGCTGTAAGTGTACGTTGAGAAAGATTCCTGCAAGCCCGCGCTGCTGCGGTTTT 708
 Db 624 CCGCTGTAAGTGTACGTTGAGAAAGATTCCTGCAAGCCCGCGCTGCTGCGGTTTT 568
 QY 709 TCCATAGGCTCCCGCCCTCGAGCAGCATCACAATAATCGACGCTCAAGTCAGAGTGGC 768
 Db 567 TCCATAGGCTCCCGCCCTCGAGCAGCATCACAATAATCGACGCTCAAGTCAGAGTGGC 508
 QY 769 GAAACCCGACGACTATAAAGATACGAGCGTTTCCCTCGGAAGCTCCCTCGTGGCT 828
 Db 507 GAAACCCGACGACTATAAAGATACGAGCGTTTCCCTCGGAAGCTCCCTCGTGGCT 448
 QY 829 CTCCTGTTCCGACCTGCGCTTACCGGATACCTGTCCGCTTTTCTCCCTTCGGGAAGCG 888
 Db 447 CTCCTGTTCCGACCTGCGCTTACCGGATACCTGTCCGCTTTTCTCCCTTCGGGAAGCG 388
 QY 889 TGGCGCTTTCTCAATGCTACGCTGAGGTATCTCAGTTCCGCTGAGTTCGCTGCCA 948
 Db 387 TGGCGCTTTCTCATAGCTACGCTGAGGTATCTCAGTTCCGCTGAGTTCGCTGCCA 328

```

QY 949 AGCTGGCGCTGTCAGCAACCCCGCTTACAGCCCGACCGCTCGGCTTATCCGGTAACCT 1008
Db 327 AGCTGGCGCTGTCAGCAACCCCGCTTACAGCCCGACCGCTCGGCTTATCCGGTAACCT 268
QY 1009 ATCGTCTTGAGTCCAAACCCGGTAAGACACGACTTATCGCCACTGGCAGCAGCCACTGGTA 1068
Db 267 ATCGTCTTGAGTCCAAACCCGGTAAGACACGACTTATCGCCACTGGCAGCAGCCACTGGTA 208
QY 1069 ACAGGATTAGCAGGAGGAGTATGTAGGGGCTGCTACAGAGTCTTTGAAGTGGTGCCTTA 1128
Db 207 ACAGGATTAGCAGGAGGAGTATGTAGGGGCTGCTACAGAGTCTTTGAAGTGGTGCCTTA 148
QY 1129 ACTACGGCTACACTAGAAGACAGTATTTGGTATCTGCGCTCTGCTGAAGCCAGTTTACCT 1188
Db 147 ACTACGGCTACACTAGAAGACAGTATTTGGTATCTGCGCTCTGCTGAAGCCAGTTT-OCT 89
QY 1189 TCGGAAAAAGAGTTGCTAGCTCTTGATCCGGCAACAAACCACCGCTGTAGGGTGGTT 1248
Db 88 TCGGAAAAAGAGTTGCTAGCTCTTGATCCGGCAACAAACCACCGCTGTAGGGTGGTT 29
QY 1249 TTTTCTTTTGAAGCAGCAGATTACGC 1275
Db 28 TTTTCTTTTGAAGCAGCAGATTACGC 2

RESULT 5
ID AAT38557/c
XX AAT38557 standard; DNA; 2932 BP.
AC AAT38557;
XX
DT 21-JAN-1997 (first entry)
XX
DE Expression vector pIH3.
XX
KW CD19; cell surface antigen; T lymphocyte; immunoselection cloning;
XX vector; pIH3; cyclic.
XX Synthetic.
OS
FH Key Location/Qualifiers
FT misc_feature 1..589 /*tag= a
FT /*note= "bases derived from pMB1 origin (pBR322 ori)"
FT misc_feature 590..597
FT /*tag= b
FT /*function= SacI linker
FT misc_feature 598..799
FT /*tag= c
FT /*function= synthetic tyrosine suppressor tRNA gene
FT LTR 800..947
FT /*tag= d
FT /*function= ASV LTR fragment (PvuII to MluI)
FT enhancer 948..1500
FT /*tag= e
FT /*function= human cytomegalovirus AD169 enhancer
FT TATA_signal 1501..1650
FT /*tag= f
FT /*function= HIV TATA and tat-responsive elements
FT misc_feature 1651..1761
FT /*tag= g
FT /*function= pLUNXAN polylinker (HindIII to Xba)
FT misc_signal 1717..2569
FT /*tag= h
FT /*function= pSV splice and poly-Addition signals
FT rep_origin 2570..2917
FT /*tag= i
FT /*function= SV40 origin of replication (PvuII to HindIII)
FT misc_feature 2918..2922
FT /*tag= j
FT

```

```

FT XX /function= pIVX, remnant of R1 site from polylinker
PN EP739980-A2.
XX
PD 30-OCT-1996.
XX
PF 23-FEB-1989; 89EP-0103127.
XX
PR 25-FEB-1988; 88US-0160416.
XX
PA (MASS-) MASSACHUSETTS GEN HOSPITAL.
XX
PI Allen J, Aruffo A, Camerini D, Lauffer L, Oquendo C;
PI Seed B, Simmons D, Stamenkovic I, Stengelin S;
XX WPI; 1996-478744/48.
DR
XX
XX cDNA encoding CD19 cell surface antigen - useful for recombinant
PT antigen prodn. for diagnostic and therapeutic purposes
XX
PS Disclosure; Fig 1; 23pp; English.
XX
XX Expression vector pIH3 (AAT38557), deposited as ATCC 67634, can be
CC used in a novel immunoselection cloning method for cell surface
CC antigen genes, e.g. human CD19 (see also AAT38556). This involves
CC introducing cDNA fragments into the vector, transfecting mammalian
CC cells in culture, maintaining the cells under conditions that allow
CC antigen expression, exposing the cells to an antibody directed
CC against the antigen, and recovering positive cells by contact with
CC an immobilised antibody directed against the first antibody.
XX
SQ Sequence 2932 BP; 773 A; 700 C; 682 G; 777 T; 0 other;

Query Match 48.8%; Score 755; DB 17; Length 2932;
Best Local Similarity 97.6%; Pred. No. 3.9e-219;
Matches 788; Conservative 0; Mismatches 15; Indels 4; Gaps 2;

QY 469 ATCGAATCTTTTCGGACTTTTGAAGTGTATGTTGGGGGAAGGATTCGAACCTTCGAA 528
Db 804 AGCAGATCTTTTCGGACTTTTGAAGTGTATGTTGGGGGAAGGATTCGAACCTTCGAA 745
QY 529 GTCGATGACGCGCAGATTAGAGTCTCTCCCTTTGGCGCTCGGGAACCCACACCGGT 588
Db 744 GTCGATGACGCGCAGATTAGAGTCTCTCCCTTTGGCGCTCGGGAACCCACACCGGT 685
QY 589 AATGCTTTTACTGGCTGCTCCCTTATCGGGAAGGGGGCGCATCATCAATGACGCG 648
Db 684 AATGCTTTTACTGGCTGCTCCCTTATCGGGAAGGGGGCGCATCATCAATGACGCG 625
QY 649 CCGCTGTAAGTGTAGCTTTGAGAAAGAAATTCCTTCAGCGCGCGGTGCTGCGGTTT 708
Db 624 CCGCTGTAAGTGTAGCTTTGAGAAAGAAATTCCTTCAGCGCGGTGCTGCGGTTT 568
QY 709 TCCATAGGCTCCGCCCGCTGACGAGCATCACAAAATTCAGCGCTCAAGTCAGAGTGGC 768
Db 567 TCCATAGGCTCCGCCCGCTGACGAGCATCACAAAATTCAGCGCTCAAGTCAGAGTGGC 508
QY 769 GAAACCCGACAGGACTATAAAGATACCAAGGTTTCCCGCTGGAAGCTCCCTCTGCGCT 828
Db 507 GAAACCCGACAGGACTATAAAGATACCAAGGTTTCCCGCTGGAAGCTCCCTCTGCGCT 448
QY 829 CTCCTGTTCGACCGCTCCCGCTTACCGGATACCTGTCCGCTTTCTCCCTTCGGGAAGCG 888
Db 447 CTCCTGTTCGACCGCTCCCGCTTACCGGATACCTGTCCGCTTTCTCCCTTCGGGAAGCG 388
QY 889 TGGCGCTTTCTCAATGCTCAGCTGATAGTATCTCAGTTCGGTGTAGGTGCTGCTCCA 948
Db 387 TGGCGCTTTCTCAATGCTCAGCTGATAGTATCTCAGTTCGGTGTAGGTGCTGCTCCA 328
QY 949 AGCTGGCGCTGTGTCAGCAACCCCGCTTACAGCCCGACCGCTGTAGGGTAACT 1008
Db 327 AGCTGGCGCTGTGTCAGCAACCCCGCTTACAGCCCGACCGCTGTAGGGTAACT 268

```

```
QY 1009 ATCGTCTTGAGTCCAAACCCGGTAAAGACAGACTTATCGCCACTGGCAGCAGCCACTGGTA 1068
Db 267 ATCGTCTTGAGTCCAAACCCGGTAAAGACAGACTTATCGCCACTGGCAGCAGCCACTGGTA 208
QY 1069 ACAGGATTAGCAGAGCGAGGTATGTAGCGGTGCTACAGAGTTCTTTGAAGTGGTGGCCTA 1128
Db 207 ACAGGATTAGCAGAGCGAGGTATGTAGCGGTGCTACAGAGTTCTTTGAAGTGGTGGCCTA 148
QY 1129 ACTACGCTACACTAGAACGACAGTATTTGGTATCTCGCGCTCTGCTGAAGCCAGTTACCT 1188
Db 147 ACTACGCTACACTAGAACGACAGTATTTGGTATCTCGCGCTCTGCTGAAGCCAGTT-CCT 89
QY 1189 TCGGAAAGAGTTGGTAGTCTTTGATCCGGCAAAACAAACACCGCTGGTAGCGGTGGTT 1248
Db 88 TCGGAAAGAGTTGGTAGTCTTTGATCCGGCAAAACAAACACCGCTGGTAGCGGTGGTT 29
QY 1249 TTTTGTGTTGCAAGCAGCAGATTACGC 1275
Db 28 TTTTGTGTTGCAAGCAGCAGATTACGC 2

RESULT 6
AAT14702/c
ID AAT14702 standard; DNA; 2932 BP.
XX
AC
XX
XX
DT 30-OCT-1996 (first entry)
XX
DE Vector plasmid pIH3.
XX
DE Cell surface antigen; cloning; immunoselection; immunotherapy;
KW therapy; diagnosis; vector; pIH3; CD2; T-cell receptor; COS;
KW T-lymphocyte; ss.
OS Synthetic.
XX
FH Key Location/Qualifiers
FT misc_feature 1..589 a
FT /tag= a
FT /note= "bases 1-589 are derived from the pMB1
FT 590..597 origin (pBR322 ori)."
FT /tag= b
FT /note= "bases 590-597 are derived from the
FT 598..799 SacII linker."
FT misc_feature 598..799
FT /tag= c
FT /note= "bases 598-799 are derived from the
FT synthetic tyrosine suppressor tRNA
FT gene (supF gene)."
FT LTR 800..947
FT /tag= d
FT /note= "bases 800-947 are derived from a remnant
FT of the ASV LTR fragment (PvuII-MluI)."
FT enhancer 948..1500
FT /tag= e
FT /note= "bases 948-1500 are derived from the hCMV
FT AD169 enhancer."
FT misc_signal 1501..1650
FT /tag= f
FT /note= "bases 1501-1650 are derived from HIV TATA
FT and tat-responsive elements"
FT misc_feature 1651..1716
FT /tag= g
FT /note= "bases 1651-1716 are derived from the
FT pLXAN polylinker (HindIII-Xba)"
FT misc_signal 1717..2569
FT /tag= h
FT /note= "bases 1717-2569 are derived from pSV
FT to splice and poly-Addition signals"
FT misc_feature 2570..2917
```

```
FT /*tag= i
FT /note= "bases 2570-2917 are derived from the SV40
FT origin of replication (PvuII-HindIII)."
FT misc_feature 2918..2922
FT /tag= j
FT /note= "bases 2918-2922 are derived from pIVX,
FT remnant of RI site from polylinker"
XX
XX US5506126-A.
XX
XX 09-APR-1996.
XX
XX 25-FEB-1988; 88US-0160416.
XX
XX 01-DEC-1992; 92US-0983647.
XX 25-FEB-1988; 88US-0160416.
XX 13-JUL-1989; 89US-0379076.
XX 13-JUL-1990; 90US-0553759.
XX 18-OCT-1993; 93US-0139273.
XX
XX (GEHO ) GEN HOSPITAL CORP.
XX
XX Aruffo A, Seed B;
XX
XX WPI; 1996-200279/20.
XX
XX Cloning of cDNA encoding cell surface antigen - useful for isolation
XX of diagnostic and therapeutic proteins
XX
XX Example 1; Fig 1A-B; 79pp; English.
XX
XX Vector pIH3 (AAT14702) is a COS cell expression vector constructed
XX from pISV by inserting a synthetic transcription unit between the
XX suppressor tRNA gene and the SV40 origin. This novel vector allows
XX the generation of very large mammalian expression libraries and
XX yields large amounts of protein in mammalian host cells, resulting
XX in efficient selection. Efficiency of library construction is
XX comparable to that achieved with phage lambda, but the clones
XX generated are easier to manipulate. A novel immunoselection
XX cloning method was used to clone genes (see also AAT14703-04 and
XX AAT14706-26) for cell surface antigens of human lymphocytes.
XX
XX Sequence 2932 BP; 774 A; 698 C; 683 G; 777 T; 0 other;
```

```
Query Match 48.88; Score 755; DB 17; Length 2932;
Best Local Similarity 97.6%; Pred. No. 3.9e-219;
Matches 788; Conservative 0; Mismatches 15; Indels 4; Gaps 2;

QY 469 ATCGAATTCTTCGGACTTTTGAAGTGATGGTGGGGAAGGATTCGAACCTTCGAA 528
Db 804 AGCAGATTCTTCGGACTTTTGAAGTGATGGTGGGGAAGGATTCGAACCTTCGAA 745
QY 529 GTCGATGACGCGCAGATTTAGAGTCTGCTCCCTTTGGCGGCTCGGGAACCCACCACCGGT 588
Db 744 GTCGATGACGCGCAGATTTAGAGTCTGCTCCCTTTGGCGGCTCGGGAACCCACCACCGGT 685
QY 589 AATGCTTTTACTGGCCTGCTCCCTTATCGGGAAGCGGGGCGCATCATATCAAAATGACGCG 648
Db 684 AATGCTTTTACTGGCCTGCTCCCTTATCGGGAAGCGGGGCGCATCATATCAAAATGACGCG 625
QY 649 CCGCTGTAAGTGTAGCTTGAGAAAGAAATTCCTGCAGCCCGCGCTTGTGCGGTTTT 708
Db 624 CCGCTGTAAGTGTAGCTTGAGAAAGAAATTCCTGCAGCCCGCGCTTGTGCGGTTTT 568
QY 709 TCCATAGGCTCCGCCCCCTGACGAGCATCACAAAATTCAGCTCAAGTCAGAGGTGGC 768
Db 567 TCCATAGGCTCCGCCCCCTGACGAGCATCACAAAATTCAGCTCAAGTCAGAGGTGGC 508
QY 769 GAAACCCGACAGGACTATAAAGATACCAGGCTTTCGCCCTGGAAGCTCCCTCGTGCCT 828
Db 507 GAAACCCGACAGGACTATAAAGATACCAGGCTTTCGCCCTGGAAGCTCCCTCGTGCCT 448
```

```
QY 829 CTCCTGTTCCGACCTCGCGCTTACCGGATACCTGTCCGCTTCTCCCTTCGGGAACG 888
Db 447 CTCCTGTTCCGACCTCGCGCTTACCGGATACCTGTCCGCTTCTCCCTTCGGGAACG 388
QY 889 TGGGCTTTCTCAATGCTCAGCTGAGGTATCTCAGTTCCGTTGAGTGTTCGCTCCA 948
Db 387 TGGGCTTTCTCAATGCTCAGCTGAGGTATCTCAGTTCCGTTGAGTGTTCGCTCCA 328
QY 949 AGCTGGCTGTGTGACGAAACCCCGCTTTCAGCCCGACCGCTTATCCGGTAACT 1008
Db 327 AGCTGGCTGTGTGACGAAACCCCGCTTTCAGCCCGACCGCTTATCCGGTAACT 268
QY 1009 ATCGTCTTGAGTCCAAACCCCGGTGAAGACACGACTTATGCCACTGGCAGCCACTGGTA 1068
Db 267 ATCGTCTTGAGTCCAAACCCCGGTGAAGACACGACTTATGCCACTGGCAGCCACTGGTA 208
QY 1069 ACAGGATTAGCAGCGAGGTATGTAGCGGTGCTACAGAGTTCTTCAAGTGTGGCCTA 1128
Db 207 ACAGGATTAGCAGCGAGGTATGTAGCGGTGCTACAGAGTTCTTCAAGTGTGGCCTA 148
QY 1129 ACTACGCTACACTAGAAGACAGTATTTGGTATCTCGCTCTGCTCAAGCCAGTTACCT 1188
Db 147 ACTACGCTACACTAGAAGACAGTATTTGGTATCTCGCTCTGCTCAAGCCAGTT-CTT 89
QY 1189 TCGGAAAGAGTTGGTACCTCTTGATCCGGGAAACCAACCCAGCTGTGTAGCGGTGTT 1248
Db 88 TCGGAAAGAGTTGGTACCTCTTGATCCGGGAAACCAACCCAGCTGTGTAGCGGTGTT 29
QY 1249 TTTTGTGTTGCAAGCAGCAGATTACGC 1275
Db 28 TTTTGTGTTGCAAGCAGCAGATTACGC 2
```

RESULT 7

AAV63441/C

XX AAV63441 standard; DNA; 2932 BP.

AC AAV63441;

XX 07-JUN-1999 (first entry)

DE Expression vector pIH3.

KW Vector; pIH3; cell surface antigen; lymphocyte; human; cloning;

KW COS; ss...

XX Chimeric - Escherichia coli.

OS Chimeric - Human cytomegalovirus.

OS Chimeric - Human immunodeficiency virus.

XX Chimeric - Rhesus macaque polyoma virus.

FH Key Location/Qualifiers

FT misc_feature 1..589

FT /*tag= a

FT /*note= "from pMB1 origin (pBR322 ori)"

FT misc_feature 590..597

FT /*tag= b

FT /*note= "Sall linker"

FT misc_feature 598..799

FT /*tag= c

FT /*note= "from synthetic tyrosine suppressor tRNA

FT LTR 800..947

FT /*tag= d

FT /*note= "ASV LTR fragment"

FT enhancer 948..1500

FT /*tag= e

FT /*note= "human cytomegalovirus AD169 enhancer"

FT TATA_signal 1501..1650

FT /*tag= f

FT /*note= "HIV TATA and tat-responsive elements"

FT misc_feature 1651..1716

FT /*tag= g

FT misc_signal /note= "pILNANXAN polylinker"

FT 1717..2569

FT /*tag= h

FT /*note= "SV40 small t antigen splice and early

FT region polyA signals from pSV2"

FT 2570..2917

FT /*tag= i

FT /*note= "SV40 origin of replication"

FT 2918..2922

FT /*tag= j

FT /*note= "from pIVX, remnant of RI site from

FT polylinker"

US5830731-A.

PD 03-NOV-1998.

XX 21-MAY-1997; 97US-0861205.

XX 01-DEC-1992; 92US-0983647.

PR 25-FEB-1988; 88US-0160416.

PR 13-JUL-1989; 89US-0379076.

PR 13-JUL-1990; 90US-0553759.

PR 21-MAY-1997; 97US-0861205.

XX (GEO) GEN HOSPITAL CORP.

XX Aruffo A, Seed B;

XX WPI; 1998-609251/51.

DR New cloning vector and polylinker - based on existing sequences for

PT efficient cloning and expression of mammalian cDNA(s), especially

PT human lymphocyte antigenic sequences

PS Example 1; Fig 1A-B; 75pp; English.

XX This is the nucleotide sequence of expression vector pIH3, a COS

CC cell expression vector constructed from pISV by inserting a

CC synthetic transcription unit between the suppressor tRNA gene and

CC the SV40 origin. The vector can be used in a novel method for

CC cloning cDNAs from mammalian expression libraries. The method is

CC based on transient expression of an antigen in eukaryotic cells and

CC immunoselection of cells expressing the antigen by adhesion to

CC an antibody-coated substrate. The method is useful for the

CC isolation and molecular cloning of any protein which can be

CC expressed and transported to the cell surface membrane of a

CC eukaryotic cell. It has been used to clone genes encoding cell

CC surface antigens associated with mammalian T lymphocyte antigens

CC such as CD1-53, LFA-3, ICAM-1, FCRI, TLISA and Leu8 (see AAV63442-63

CC and AAW80440-55). The invention provides high efficiency expression

CC vectors, such as pIH3, which allow the generation of very large

CC mammalian expression libraries, and yield large amounts of protein

CC in mammalian host cells, resulting in efficient selection. The

CC purified genes and proteins are useful for immunodiagnostic and

CC immunotherapeutic applications, including the diagnosis and

CC treatment of immune-mediated infections, diseases, and disorders of

CC animals, including humans.

SQ Sequence 2932 BP; 774 A; 699 C; 682 G; 777 T; 0 other;

Query Match 48.8%; Score 755; DB 19; Length 2932;

Best Local Similarity 97.6%; Pred. No. 3.9e-219;

Matches 788; Conservative 0; Mismatches 15; Indels 4; Gaps 2;

QY 469 ATCGAATCTTCGGACCTTTTGAAGTGTGTTGGGGGAAGGATTGGAACCTTCGAA 528

Db 804 AGCAGATTCTTTTCGGACCTTTTGAAGTGTGTTGGGGGAAGGATTGGAACCTTCGAA 745

QY 529 GTCGATGACGGCAGATTATAGAGTCTGCCCTTTGGCGGCTCGGGAACCCACACGGGT 588

Db 744 GTCGATGACGGCAGATTATAGAGTCTGCCCTTTGGCGGCTCGGGAACCCACACGGGT 685

Query Match 48.8%; Score 755; DB 20; Length 2932;
 Best Local Similarity 97.6%; Pred. No. 3.9e-219;
 Matches 788; Conservative 0; Mismatches 15; Indels 4; Gaps 2;

QY 469 ATCGAATCTCTTCGGACTTTTGAAGATGATGTGGTGGGGAAGATTGCAACCTTCGAA 528
 DB 804 AGCAGATCTTTTCGGACTTTTGAAGATGATGTGGTGGGGAAGATTGCAACCTTCGAA 745

QY 529 GTCGATGAGCGCAGATTTAGAGTCGTCTCCCTTTGCGCGCTCGGGAACCCACACCGGT 588
 DB 744 GTCGATGAGCGCAGATTTAGAGTCGTCTCCCTTTGCGCGCTCGGGAACCCACACCGGT 685

QY 589 AATGCTTTTACTGCGCTCTGCTTATCGGAAGCGGCGCATCATATCAATGACGCG 648
 DB 684 AATGCTTTTACTGCGCTCTGCTTATCGGAAGCGGCGCATCATATCAATGACGCG 625

QY 649 CGGCTGTAAAGTGTACCTTGAGAAAGATTCCTGCAGCCCGCGCTGCTGCGGTTT 708
 DB 624 CGGCTGTAAAGTGTACCTTGAGAAAGATTCCTGCAGCCCGCGCTGCTGCGGTTT 568

QY 709 TCCATAGGCTCGGCGCCCTGACGAGCATCAAAAAATCGACGCTCAAGTCAGAGTGGC 768
 DB 567 TCCATAGGCTCGGCGCCCTGACGAGCATCAAAAAATCGACGCTCAAGTCAGAGTGGC 508

QY 769 GAAACCCGACAGGACTATAAGATACAGCGCTTCCCTCGAAGCTCCCTCGTGGCT 828
 DB 507 GAAACCCGACAGGACTATAAGATACAGCGCTTCCCTCGAAGCTCCCTCGTGGCT 448

QY 829 CTCCTGTTCCGACCTCGCGTTACCGGATACCTGTCGCGCTTCTCCCTTCGGGAGCG 888
 DB 447 CTCCTGTTCCGACCTCGCGTTACCGGATACCTGTCGCGCTTCTCCCTTCGGGAGCG 388

QY 889 TGGCGCTTCTCAATGCTCAGCTGTAGGTATCTCAGTTCGGGTAGGTTCGCTCCA 948
 DB 387 TGGCGCTTCTCAATGCTCAGCTGTAGGTATCTCAGTTCGGGTAGGTTCGCTCCA 328

QY 949 AGTGGGCTGTGTGACGAAACCCCGCTTCAGCCGACGCTCGCTTATCCGTTAACT 1008
 DB 327 AGTGGGCTGTGTGACGAAACCCCGCTTCAGCCGACGCTCGCTTATCCGTTAACT 268

QY 1009 ATCGTCTGAGTCCACCCGGTAAGACAGCTTATCGCCACTGGCAGCAGCTGTA 1068
 DB 267 ATCGTCTGAGTCCACCCGGTAAGACAGCTTATCGCCACTGGCAGCAGCTGTA 208

QY 1069 ACAGGATTAGCAGCGAGGTATGTAGCGGTGCTACAGCTTGTGAAGTGTGGCCTA 1128
 DB 207 ACAGGATTAGCAGCGAGGTATGTAGCGGTGCTACAGCTTGTGAAGTGTGGCCTA 148

QY 1129 ACTACGCTACACTAGAGCAGATTTGGTATCTGCGCTCTGCTGAAGCCAGTTACT 1188
 DB 147 ACTACGCTACACTAGAGCAGATTTGGTATCTGCGCTCTGCTGAAGCCAGTTACT 89

QY 1189 TCGGAAAAAGTTGTAGCTCTTATCGGGAACCAACACCTGCTGCTGAGCGTGGT 1248
 DB 88 TCGGAAAAAGTTGTAGCTCTTATCGGGAACCAACACCTGCTGCTGAGCGTGGT 29

QY 1249 TTTTGTGTTGCAAGCAGCAGATTACGC 1275
 DB 28 TTTTGTGTTGCAAGCAGCAGATTACGC 2

RESULT 9
 AAA50577/c

ID AAA50577 standard; DNA; 2932 BP.

XX AAA50577;

XX 19-DEC-2000 (first entry)

DT Eukaryotic cell expression vector plasmid piH3.

DE Plasmid piH3; vector; eukaryotic cell; COS; cell surface antigen;

XX

KW

immunodiagnosis; diagnosis; immunotherapy; gene therapy;
 immune disorder; infection; asthma; immune-complex disease;
 anyloidosis; multiple sclerosis; parasitic disease;
 immunoselection; panning; ds.

OS Chimeric - Rhesus macaque polyoma virus.
 OS Chimeric - Escherichia coli.
 OS Chimeric - Human cytomegalovirus.
 OS Chimeric - Human immunodeficiency virus.
 OS Chimeric - Synthetic.

XX Key Location/Qualifiers
 FT misc_feature 1..589
 FT /tag= a
 FT /function= "replication origin"
 FT /note= "pBR322 ori from pMB1"
 FT misc_feature 590..597
 FT /tag= b
 FT /note= "derived from SacI linker"
 FT misc_feature 598..799
 FT /tag= c
 FT /gene= "supF"
 FT /note= "synthetic tyrosine suppressor tRNA gene"
 FT LTR 800..947
 FT /tag= d
 FT /note= "remnant of ASV LTR fragment"
 FT enhancer 948..1500
 FT /tag= e
 FT /note= "human cytomegalovirus AD169 enhancer"
 FT TATA_signal 1501..1650
 FT /tag= f
 FT /note= "HIV TATA and tat-responsive element"
 FT misc_feature 1651..1716
 FT /tag= g
 FT /note= "derived from pILNAN polylinker"
 FT polyA_site 1717..2569
 FT /tag= h
 FT /note= "derived from PSV to splice and poly-A site"
 FT misc_feature 2570..2917
 FT /tag= i
 FT /function= "replication origin"
 FT /note= "from SV40 origin of replication"
 FT misc_feature 2918..2922
 FT /tag= j
 FT /note= "from pIVX"

US6111093-A.
 29-AUG-2000.
 28-OCT-1998; 98US-0181612.
 01-DEC-1992; 92US-0983647.
 25-FEB-1988; 88US-0160416.
 13-JUL-1989; 89US-0379076.
 23-MAR-1990; 90US-0498809.
 13-JUL-1990; 90US-0553759.
 (GEO) GEN HOSPITAL CORP.
 Stamenkovic I, Seed B;
 WPI; 2000-586382/55.
 Isolated nucleic acid molecule encoding the CD19 cell surface antigen,
 useful for immunodiagnosis and immunotherapy of immune-mediated
 infections or disorders, e.g. asthma, immune-complex disease, parasitic
 diseases -
 Example 1; Fig 1A-B; 75pp; English.
 The present sequence is that of COS cell expression vector
 plasmid piH3. The vector was constructed from pISV by inserting

CC a synthetic transcription unit between the suppressor tRNA gene and
CC the SV40 origin. The transcription unit consisted of a chimeric
CC promoter composed of human cytomegalovirus AD169 immediate early
CC enhancer sequences fused to the HIV LTR -67 to +80 sequences.
CC Immediately downstream from the LTR +80 was inserted a polylinker
CC containing 2 BstXI sites separated by a 350 bp stuffer. The BstXI
CC sites were flanked by XbaI sites, which could also be used to
CC excise the insert. Downstream from the polylinker were placed the
CC SV40 small t antigen splice and early region polyA signals derived
CC from pSV2. The vector can be used to construct large mammalian
CC cDNA expression libraries. A new method for cloning cDNA encoding
CC cell surface antigens is based upon transient expression of antigen
CC in eukaryotic (e.g. COS) cells and physical selection of cells
CC expressing the antigen by adhesion to (panning on) an antibody
CC coated substrate such as a culture dish. The purified cDNA and the
CC protein that it encodes are useful for immunodiagnostic and
CC immunotherapeutic applications, including the diagnosis and
CC treatment of immune-mediated infections, diseases, and disorders in
CC animals, including humans. These disorders include asthma,
CC immune-complex disease, amyloidosis, parasitic diseases or multiple
CC sclerosis. High efficiency cDNA expression vectors such as pIH3
CC were designed for use in the novel methods of the invention,
CC yielding large amounts of protein in mammalian host cells,
CC resulting in efficient selection.

XX
SQ Sequence 2932 BP; 774 A; 699 C; 682 G; 777 T; 0 other;

Query Match 48.8%; Score 755; DB 21; Length 2932;
Best Local Similarity 97.6%; Pred. No. 3.9e-219;
Matches 788; Conservative 0; Mismatches 15; Indels 4; Gaps 2;

QY 469 ATCGAATCTTTCGGACTTTTGAAGTGATGGTGGGGGAAGGATTCGAACCTTCGAA 528
DB 804 ACAGATTCTTCGGACTTTTGAAGTGATGGTGGGGGAAGGATTCGAACCTTCGAA 745
QY 529 GTCGATGACGGCAGATTAGAGTCTGCTCCCTTTGGCCGCTCGGGAAACCCACACGGGT 588
DB 744 GTCGATGACGGCAGATTAGAGTCTGCTCCCTTTGGCCGCTCGGGAAACCCACACGGGT 685
QY 589 AATGCTTTTACTGGCTGCTCCCTTATCGGGAAGCGGGGCGCATCATCAATGACGG 648
DB 684 AATGCTTTTACTGGCTGCTCCCTTATCGGGAAGCGGGGCGCATCATCAATGACGG 625
QY 649 CCCTGTAAAGTGTAGCTTTGAGAAGAAATTCCTGCACGCCGCGCTCTGCGCGTTT 708
DB 624 CCCTGTAAAGTGTAGCTTTGAGAAGAA---CCGCGTAATTCGCGCTTCTGCGCGTTT 568
QY 709 TCCATAGGCTCCGCCGCCCTGACGAGCATCACAAAATCGACGCTCAAGTCAGAGGTGGC 768
DB 567 TCCATAGGCTCCGCCGCCCTGACGAGCATCACAAAATCGACGCTCAAGTCAGAGGTGGC 508
QY 769 GAAACCGGACGACTATAAGATACAGCGCTTCCCTCGAAGCTCCTCGTGGCGT 828
DB 507 GAAACCGGACGACTATAAGATACAGCGCTTCCCTCGAAGCTCCTCGTGGCGT 448
QY 829 CTCCTGTCCGACCTGCGCTTACCGGATACCTGTCGCGCTTCTCCCTTCGGGAAGCG 888
DB 447 CTCCTGTCCGACCTGCGCTTACCGGATACCTGTCGCGCTTCTCCCTTCGGGAAGCG 388
QY 889 TGGCGCTTCTCAATGCTCAGCTGAGGTATCTCAGTTCGGGTAGGTGCTGCGTCCA 948
DB 387 TGGCGCTTCTCAATGCTCAGCTGAGGTATCTCAGTTCGGGTAGGTGCTGCGTCCA 328
QY 949 AGCTGGGCTGTGTCAGCAACCCCGCTTTCAGCCCGACCGCTTATCCGGTAAC 1008
DB 327 AGCTGGGCTGTGTCAGCAACCCCGCTTTCAGCCCGACCGCTTATCCGGTAAC 268
QY 1009 ATCGCTTTGAGTCCAAACCCGGTGAAGACCACTTATCGCCACTGGCAGCAGCACTGGTA 1068
DB 267 ATCGCTTTGAGTCCAAACCCGGTGAAGACCACTTATCGCCACTGGCAGCAGCACTGGTA 208
QY 1069 ACAGGATTAGCAGCGAGGTATGTAGGGCGGTGCTACAGAGTTCTTGAAGTGGTGGCCTA 1128

DB 207 ACAGGATTAGCAGCGAGGTATGTAGCGGTGTACAGAGTTCTTGAAGTGGTGGCCTA 148
QY 1129 ACTACGGCTACACTAGAGGACAGATTTTGGTATCTGCGTCTCTGAAGCCAGTTCCT 1188
DB 147 ACTACGGCTACACTAGAGGACAGATTTTGGTATCTGCGTCTCTGAAGCCAGTTCCT 89
QY 1189 TCGGAAAAAGAGTTGGTAGCTCTTTGATCCGGCAACAAACACCCGCTGCTAGCGGTGTT 1248
DB 88 TCGGAAAAAGAGTTGGTAGCTCTTTGATCCGGCAACAAACACCCGCTGCTAGCGGTGTT 29
QY 1249 TTTTGTGTCACGACGACAGATTACGC 1275
DB 28 TTTTGTGTCACGACGACAGATTACGC 2
RESULT 10
AAS03171/c
ID AAS03171 standard; DNA; 2932 BP.
XX
AC AAS03171;
XX
DT 29-AUG-2001 (first entry)
XX
DE Expression vector pIH3 DNA used to express human lymphocyte antigens.
KW Human; Lymphocyte cell surface antigen; immune-mediated disease; asthma;
KW infection; immune deficiency disorder; hypersensitivity; inflammation;
KW systemic lupus erythematosus; platelet disorder; rheumatoid arthritis;
KW transplant rejection; cyclic; circular; mutant; piH3; ds.
XX Chimeric - Avian sarcoma virus.
OS Chimeric - Human cytomegalovirus.
OS Chimeric - Human immunodeficiency virus.
OS Chimeric - Rhesus macaque polyoma virus.
OS Synthetic.
XX
FH Key Location/Qualifiers
FT misc_feature 1..589 /*tag= a
FT /*note= "Derived from the pMB1 origin (pBR322 ori)"
FT misc_feature 590..597 /*tag= b
FT /*note= "Derived from the SacII linker (ACCGGCT)"
FT misc_feature 598..799 /*tag= c
FT /*note= "Derived from the synthetic tRNA gene (supP)"
FT misc_feature 800..947 /*tag= d
FT /*note= "Derived from a remnant of the ASV LTR
FT fragment (PvuII to MluI)"
FT enhancer 948..1500 /*tag= e
FT /*note= "Derived from the human cytomegalovirus
FT AD169 enhancer"
FT promoter 1501..1650 /*tag= f
FT /*note= "Derived from the HIV TATA and tat-responsive
FT elements"
FT misc_feature 1651..1716 /*tag= g
FT /*note= "Derived from the pILNAN polylinker
FT (HindIII to Xba)"
FT misc_signal 1717..2569 /*tag= h
FT /*note= "Derived from the pSV2 splice and polyadenylation
FT signals"
FT misc_feature 2570..2917 /*tag= i
FT /*note= "Derived from the SV40 origin of replication
FT (PvuII to HindIII)"
FT misc_feature 2918..2922 /*tag= j

XX The invention provides retroviral vectors derived from a non-primate
 CC lentivirus genome. These vectors comprise a deleted gag gene. The
 CC deletion in gag removes one or more nucleotides downstream of nucleotide
 CC 350 of the gag coding sequence. One or more accessory genes are absent
 CC from the non-primate lentivirus genome or lack the tat gene but includes
 CC the leader sequences between the end of the 5' long terminal repeat (LTR)
 CC and the ATG of gag. The vectors, particles or cells transfected with
 CC either of these, are useful for the delivery of nucleotides of interest
 CC to a target site (i.e. gene therapy). The retroviral vector can be used
 CC as a delivery system. Alternatively, a non-retroviral expression vector,
 CC adenovirus or plasmid can be used as a delivery system for the retroviral
 CC vector. The retroviral vectors are capable of transferring genetic
 CC material to non-dividing or slowly dividing cells. Deletion of larger
 CC regions of the gag gene in Equine infectious anemia virus (EIAV) vectors
 CC is advantageous and leads to higher titers of viral vector being
 CC produced. Deletion of accessory genes permits vectors to be produced
 CC without the genes normally associated with disease in lentiviral (e.g.
 CC HIV) infections. The deletion of these genes also permits the vector to
 CC package more heterologous DNA.

SQ Sequence 7252 BP; 1942 A; 1686 C; 1751 G; 1873 T; 0 other;

Query Match 39.3%; Score 607.4; DB 20; Length 7252;
 Best Local Similarity 97.5%; Pred. No. 5e-174;
 Matches 617; Conservative 0; Mismatches 16; Indels 0; Gaps 0;

QY 660 TGTACGTTGACAAAGATTCCTGACGCCGCGGCTGCTGCGCTTTTTCATAGGCTC 719
 DB 5062 tataccggtgagaacatgagactgactggcgcggtgctgctgcttttccataggtc 5121
 QY 720 CGCCGCCCTGACGAGCATCACAAAATCGACGCTCAAGTCAGAGGTGGCGAAACCCGACA 779
 DB 5122 cgcgccctgacgagcatcacaaaatgacgctcaagtcagaggtgcgaaaccgcgaca 5181
 QY 780 GACATATAAGATACACAGCGCTTCCCGCTGGAAGCTCCCTGCGGCTCTCTGTCCG 839
 DB 5182 ggactataaagatacagcggttcccccctggaagctccctcgtgctcctcgttcgcg 5241
 QY 840 ACCCTGCCGCTTACCGGATACCTGCTGCGCTTTCCTCCCTGCGGAGCGTGGCGTTCT 899
 DB 5242 accctgcgcctcaccggatcacctgctcgccttctcctctcgggaagcgtggcgcttct 5301
 QY 900 CAATGCTACGCTGTAGGTATCTCAGTTCGGTGTAGTTCGGTTCGCTCAAGCTGGGCTGT 959
 DB 5302 caatgctcaagctgtaggatctcagttcgttgggtcgttcgcccacagctgggctgt 5361
 QY 960 GTGACGAACCCCGCTTACGCCCGACCGCTGCGCTTATCCGGTAACTATCTCTTGAG 1019
 DB 5362 gtgcacgaaccccgcttcagccgctgagcgtgctcctatccggttaactatgctctgag 5421
 QY 1020 TCCAACCCGGTAAACACACATTCGCCACATTCGCCAGCGACGACCTGGTAAACAGATTAGC 1079
 DB 5422 tccaacccggtaagacacactatcgccactggcagcagcactggtaacagattagc 5481
 QY 1080 AGAGCGAGGTATGTAGCGCGTGTACAGAGTCTTCAAGTGGTGGCTTAACATACGCTAC 1139
 DB 5482 agagcgaggtatgtaggcggtgctacagaggtcttgaaagtggcgtaactacgggttac 5541
 QY 1140 ACTAGAGGACAGTATTTGGTATCTGCGCTCTGCTGAAGCCAGTTACCTTCGSAAGAGA 1199
 DB 5542 actagagggacagatttggtatctgctgctgctgctgctgctgctgctgctgctgctgct 5601
 QY 1200 GTTGTAGTCTTGTATCCGGCAACAAACACCGCTGGTAGCGGTGGTTTTTGTGTTGC 1259
 DB 5602 gttggtagctctgtatccggcaacaaacacccgctggtgagcgtggtggtgtttttgttgc 5661
 QY 1260 AAGCAGCAGATTACGCGCAGAAAAAAGGATCT 1292
 DB 5662 aagcagcagattacgcgcagaaaaaaggatct 5694

RESULT 12
 AAT90692
 ID AAT90692 standard; DNA; 5865 BP.
 XX
 AC AAT90692;
 XX
 DT 05-JAN-1998 (first entry)
 XX
 DE Plasmid FBdelPRDSAF coding sequence.
 XX
 KW Packaging-deficient construct; viral gag-pol gene; packaging cell line;
 KW moloney murine leukaemia virus; MoMLV; viral env gene; helper construct;
 KW gene therapy; ss.
 XX
 OS Synthetic.
 XX
 PN WO9708330-A1.
 XX
 PD 06-MAR-1997.
 XX
 PF 23-AUG-1996; 96WO-GB02061.
 XX
 PR 23-AUG-1995; 95GB-0017263.
 XX
 (CANC-) CANCER RES CAMPAIGN TECHNOLOGY.
 PA
 PI Collins MKL, Cosset F, Takeuchi Y, Weiss RA;
 XX
 DR WPI; 1997-179287/16.
 XX
 PT Selectable retroviral packaging cell lines and expression constructs
 PT - comprise selectable gene downstream of gene of interest, are
 PT selectable due to the in-efficiency associated with translation
 PT re-initiation
 XX
 PS Claim 13; Fig 12; 79pp; English.

XX This sequence represents the recombinant expression plasmid FBdelPASAF.
 CC This sequence is a packaging-deficient construct having a viral env gene
 CC (in this case from Rul14) and a selectable marker (SM). It is an example
 CC of a recombinant expression vector (REV) of the invention, used to create
 CC a packaging cell line. The REV's of the invention comprise a gene of
 CC interest (GOI) and a SM gene. The SM gene is arranged downstream of the
 CC GOI and a GOI associated stop codon is spaced from a start codon of the
 CC SM gene to ensure that the SM protein is expressed as a result of
 CC translation reinitiation. The cell lines are transformed with two REV's,
 CC both are replication deficient, one contains the viral gag-pol gene, the
 CC other the viral env gene. By using helper constructs, such as the REV's,
 CC which are directly selectable and which provide for high expression of
 CC the viral gene, high titre retroviral vectors may be obtained. The
 CC packaging cell lines are useful for gene therapy. Prior packaging cell
 CC lines using full length retroviral genomes as helper genomes were
 CC isolated by cotransfecting them with plasmids encoding selectable
 CC markers. However, the helper functions can be lost during the passages of
 CC the cells in culture and the current packaging systems provide limited
 CC titres of infectious retroviral vectors. Co-transfection with a plasmid
 CC encoding a SM does not directly select the best gag-pol-env-expressing
 CC cells. The new retroviral packaging cell lines overcome these problems.
 XX
 SQ Sequence 5865 BP; 1492 A; 1602 C; 1390 G; 1373 T; 8 other;

Query Match 39.1%; Score 604.4; DB 18; Length 5865;
 Best Local Similarity 91.4%; Pred. No. 3.7e-173;
 Matches 638; Conservative 0; Mismatches 60; Indels 0; Gaps 0;
 QY 595 TTCTACTGGCTGCTCCCTTATCGGGAAGCGGGCGCATCATCAATATACGCCGCTG 654
 DB 3752 ttatcatgctggtccagatctggcccatcgccgagcagatcatnnnnacatgtgag 3811
 QY 655 TAAAGTGTACGTGTGAGAAAGAAATTCCTGACGCCCGCGCTGCTGGCGTTTTCATA 714
 DB 3812 caaaagggcagcaaaagggccaggaacgcgtataaaggcgctgtgctgctgtttccata 3871

QY	715	GGCTCCGCCCCCTGAGGAGCATCAAAAATCGACGCTCAAGTCAGAGGTGGCGAAACC	774
Db	3872	ggctccgccccctgacgagcatcaaaaaatcgacgctcaagtcagaagtggcgaaacc	3931
QY	775	CGACGAGACTATAAGATACCAGGCGTTTCCCCCTGGGAAGTCCCTCGTGGCGCTCTCCTG	834
Db	3932	cgaagagactataaagataccaggctttcccccctggaagctccctcgtgcgctcctcg	3991
QY	835	TTCCGACCCCTCGCGCTTTACCGGATACCTCTCGCGCCTTCTCCCTTCGGGAAGCGTGGCGC	894
Db	3992	ttcgaacctgcgcgttacccggataccctgcgcgccttctcccttcgggaagcgtggcgc	4051
QY	895	TTTCTCAATGCTCACGCTGAGTATCTCAGTTCGGGTAGGTGGTTCGGTCCCAAGCTGG	954
Db	4052	tttctcaatgctcaecgtgtaggtatctcagttcggtaggttcgtcccaagctgg	4111
QY	955	GCTGTGTGACGAACCCCGTTCAGCCGCGCTGCGCCTTATTCGGGTAACTATGTC	1014
Db	4112	gctgtgtgcagcaacccccgttccagccgcgcgtgcgcctatccggttaactatcgtc	4171
QY	1015	TTGAGTCCAAACCCGGTAAGACAGCACTTATGCCACTGGCAGCAGCCACTGGTAACAGGA	1074
Db	4172	ttgaagccaacccggtaaagacagcacttatgcacactggcagcgcacactcggtaaacagga	4231
QY	1075	TTAGCAGAGGAGGTATGTAGCGGTGCTACAGAGTCTTTGAACTGGTGGCCTAACTACG	1134
Db	4232	ttagcagagcgaggtatgtaggcggtgtacagagttcttgaagtgtggcctaactacg	4291
QY	1135	GCTACACTAGAGGACAGTATTGGTATCTGGCTCTGCTGAAGCCAGTTACCTTCGGAA	1194
Db	4292	gctacactagaaggaagcatttggtagtctgcgtctgctgaagccaagttacacttcggaa	4351
QY	1195	AAAGAGTGGTAGCTTTGATCCGCCAAACAAACACCGCTGGTGGCGGTGGTTTTTTTG	1254
Db	4352	aaagagttggtagtcttgatccggcaaaaaaacacacgcgtggtagcgttgggttttttgg	4411
QY	1255	TTTTCGACAGCAGCATTTACGCCGAGAAAAAAGGATCT	1292
Db	4412	tttgcgaagcagcagattacgcgcagaaaaaaggatct	4449

```

RESULT 13
AAT90689
ID AAT90689 standard; DNA; 6028 BP.
XX
XX
AC AC
XX
XX
DT 05-JAN-1998 (first entry)
XX
XX Plasmid FBdelpASAF coding sequence.
XX
XX Packaging-deficient construct; viral gag-pol gene; packaging cell line;
KW moloney murine leukaemia virus; MoMLV; viral env gene; helper construct;
KW gene therapy; ss.
XX
XX Synthetic.
XX
XX WO9708330-A1.
XX
XX 06-MAR-1997.
XX
XX 23-AUG-1996; 96WO-GB02061.
XX
XX 23-AUG-1995; 95GB-0017263.
XX
XX (CANC-) CANCER RES CAMPAIGN TECHNOLOGY.
XX
XX Collins MKL, Cosset F, Takeuchi Y, Weiss RA;
PI
XX WPI; 1997-179287/16.
XX
XX Selectable retroviral packaging cell lines and expression constructs

```

- comprise selectable gene downstream of gene of interest, are selectable due to the in-efficiency associated with translation re-initiation

Claim 13; Fig 9; 79pp; English.

This sequence represents the recombinant expression plasmid FBdelPASAF. This sequence is a packaging-deficient construct having a viral env gene (in this case from the moloney murine leukaemia virus) and a selectable marker (SM). It is an example of a recombinant expression vector (REV) of the invention, used to create a packaging cell line. The REVs of the invention comprise a gene of interest (GOI) and a SM gene. The SM gene is arranged downstream of the GOI and a GOI associated stop codon is spaced from a start codon of the SM gene to ensure that the SM protein is expressed as a result of translation reinitiation. The cell lines are transformed with two REVs, both are replication deficient, one contains the viral gag-pol gene, the other the viral env gene. By using helper constructs, such as the REV's, which are directly selectable and which provide for high expression of the viral gene, high titre retroviral vectors may be obtained. The packaging cell lines are useful for gene therapy. Prior packaging cell lines using full length retroviral genomes as helper genomes were isolated by cotransfecting them with plasmids encoding selectable markers. However, the helper functions can be lost during the passages of the cells in culture and the current packaging systems provide limited titres of infectious retroviral vectors. Co-transfection with a plasmid encoding a SM does not directly select the best gag-pol-env-expressing cells. The new retroviral packaging cell lines overcome these problems.

Sequence 6028 BP; 1515 A; 1637 C; 1499 G; 1369 T; 8 other;

Query Match 39,1%; Score 604.4; DB 18; Length 6028;
Best Local Similarity 91.4%; Pred. No. 3.7e-173;
Matches 638; Conservative 0; Mismatches 60; Indels 0; Gaps 0;

Query Match	39.1%	Score 604.4	DB 18	Length 6028
Best Local Similarity	91.4%	Pred. No. 3.7e-173		
Matches 638	Conservative 0	Mismatches 60	Indels 0	Gaps 0
QY	595	TTTACTGCGCTCCCTTATCGGGAACGGGGGCATCATATCAATCAGCGCGCGCTG	654	
Db	3915	ttatcatgtctggtatccagatcttgggccatcgcgccgcatgatnnnmacatgtgag	3974	
QY	655	TAAAGTGTACGTTTGAGAAAGAAATCTCTGCAGCCGCGCTTGCTGCGGTTTTCATAT	714	
Db	3975	caaaagccagcaaaagccaggaaacgtaaaaggccgctgtcgcggttttccata	4034	
QY	715	GGCTCCGCCCCCTGCAGGAGCATCACAAAATCAGCGCTCAAGTCAGAGGTGGCGAAACC	774	
Db	4035	ggctccgccccctgacgagcatcacaaaatcagcgtcaagtcagaggtggcgaaacc	4094	
QY	775	CGACGAGCACTATAAGATACACAGCGTTTCGCCCTGGAAGCTCCCTGCGCTCCCTG	834	
Db	4095	cgacgagcataaagataccaggcgtttccccctggaaagctccctcgtcgctctcctg	4154	
QY	835	TTCGGACCTGCGCGTTACCGGATACCTGTCGCGCTTCTCCCTTCGSGAAGCTGGCGC	894	
Db	4155	ttcgcagccctgcgcttaaccggataacctgtccgcttctcccttcgggaagcgtggcg	4214	
QY	895	TTTCTCAATGCTCACGCTGTAGGTATCTCAGTTCGGTGTAGGTGCTTCGAAGCTGG	954	
Db	4215	ttctcaatgctcacgctgtaggtatctcagttcgggtgtaggtgcttcgctccaagctg	4274	
QY	955	GCTGTGTGCACGAACCCCGCTTCAGCCGACGCTGCGCTTATCCGGTAACATATCGTC	1014	
Db	4275	gctgtgtgaacgaaccccccggttcagcccgagcgcgtgccttatccggttaactatcgtc	4334	
QY	1015	TTGAGTCCAAACCGGTAAACACAGCACTTATGGCACTGGCAGCAGCCACTGGTAACAGGA	1074	
Db	4335	ttgagttccaaacccgtaagacacgaacttafcgcactggcagcagccactggtaaacagga	4394	
QY	1075	TTAGCAGAGCGAGGTATGAGGCGGTGCTACAGAGTCTTGAAGTGGTGGCCCTAACCTACG	1134	
Db	4395	ttagcagagcgaggtatgtaggcggcggtactacagagttcttgaagtggtggcccttaactac	4454	
QY	1135	GCTACTAGAAAGCACAGTATTTGGTATCTGCCCTCTGCTGGAAGCCAGTTTACCTTCGGAA	1194	

Db	3975	caaaagccagcaaaagccaggaacgcgtataaaagccgcgtctgctggcggttttcccata	4034
QY	715	GGCTCCGCCCCCTTGACGAGCATCACAAAAATCGACGCTCAAGTCAGAGGTGGCGAAACC	774
Db	4035	ggctccgcccccttgcagcagcatcacaaaatcgacgcgtcaagtcagaggtggcgaacc	4094
QY	775	CGACGAGACTATAAAGATACCAAGCGTTTCGCCCTGGAAAGTCCTCGTGCCTCTCCTG	834
Db	4095	cgacagactataaagataccagcggttccccctggaaagctccctcgcgcgtctccgtg	4154
QY	835	TTCCGACCCTGCGCCTTACCGGATACCTGTGCCGCTTTCCTCCCTCGGGAAGCGTGCGC	894
Db	4155	ttccgacctgcgccttaccggataacctgtccgccttctcccttcgggaagcgtggcgc	4214
QY	895	TTTCTCAATGCTCACGCTGTAGGTATCTCAGTTCGGGTGTAGTGTCTGCCTCCAAAGTGG	954
Db	4215	ttctcaatgctcacgctglagtatctcagtttcggtgtagggctgttcgctcccaagctgg	4274
QY	955	GCNTGTGACGAAACCCCGTTACGCCCGACGCTGGCCTTATCGCGTAACATATCGTC	1014
Db	4275	gcctgtgcacgaaccccccggttcagccgaccgctgccttatccggaatactcgtc	4334
QY	1015	TTGAGTCCAAACCCGGTAAGACAGGACTATTCGCCACTGGCAGCAGCCACTGGTAACAGGA	1074
Db	4335	ttgagtcacaaccgctaagacacgacttatcgccactggcagcagccaclggtaaacagga	4394
QY	1075	TTACGACAGCGAGGTATGTAGCGGTGCTACAGAGTCTTTGAAGTGGTGGCTTAACTAGC	1134
Db	4395	ttacgcagcgcaggtatgtaggcgggctacagagttcttgaagtggtgacctaacctacg	4454
QY	1135	GCTACACTAGAAGGACAGTATTTGGTATCTCGCCTCTGCTGAAGCCAGTTTACCTTCGGAA	1194

||||| gctacactagaagacagtatttggtatctgcgtctgctgaagccagttacacctcgaa 4514
Db 4455
QY 1195 AAGAGTTGCTAGCTCTTGATCCGGCAACAAACACCGCTGCTAGCGGTGTTTGTG 1254
Db 4515 aagagttgtagctctgtatccgcgaacaaacacccgcgtgtagcgttggtttttg 4574
QY 1255 TTGCAAGCAGCAGATTACCGCGAGAAAAAAGGATCT 1292
Db 4575 ttgcaagcagcagattacgcgcagaaaaaaggatct 4612

RESULT 14
ID AAT90690 standard; DNA; 6061 BP.
XX
AC AAT90690;
XX
DT 05-JAN-1998 (first entry)
XX
DE Plasmid FBdelPMOSAF coding sequence.
XX
KW Packaging-deficient construct; viral gag-pol gene; packaging cell line;
KW moloney murine leukaemia virus; MOMLV; viral env gene; helper construct;
KW gene therapy; ss.
XX
OS Synthetic.
XX
PN WO9708330-A1.
XX
PD 06-MAR-1997.
XX
XX 23-AUG-1996; 96WO-GB02061.
XX
PR 23-AUG-1995; 95GB-0017263.
XX
PA (CANC-) CANCER RES CAMPAIGN TECHNOLOGY.
XX
XX Collins MKL, Cosset F, Takeuchi Y, Weiss RA;
XX WPI; 1997-179287/16.
XX
XX Selectable retroviral packaging cell lines and expression constructs
PT - comprise selectable gene downstream of gene of interest, are
PT selectable due to the in-efficiency associated with translation
PT re-initiation
XX
PS Claim 13; Fig 10; 79pp; English.
XX
CC This sequence represents the recombinant expression plasmid FBdelPASAF.
CC This sequence is a packaging-deficient construct having a viral env gene
CC (in this case from the moloney murine leukaemia virus) and a selectable
CC marker (SM). It is an example of a recombinant expression vector (REV) of
CC the invention, used to create a packaging cell line. The REVS of the
CC invention comprise a gene of interest (GOI) and a SM gene. The SM gene is
CC arranged downstream of the GOI and a GOI associated stop codon is spaced
CC from a start codon of the SM gene to ensure that the SM protein is
CC expressed as a result of translation reinitiation. The cell lines are
CC transformed with two REVS, both are replication deficient, one contains
CC the viral gag-pol gene, the other the viral env gene. By using helper
CC constructs, such as the REV's, which are directly selectable and which
CC provide for high expression of the viral gene, high titre retroviral
CC vectors may be obtained. The packaging cell lines are useful for gene
CC therapy. Prior packaging cell lines using full length retroviral genomes
CC as helper genomes were isolated by cotransfecting them with plasmids
CC encoding selectable markers. However, the helper functions can be lost
CC during the passages of the cells in culture and the current packaging
CC systems provide limited titres of infectious retroviral vectors.
CC Co-transfection with a plasmid encoding a SM does not directly select the
CC best gag-pol-env-expressing cells. The new retroviral packaging cell
CC lines overcome these problems.
XX
XX Sequence 6061 BP; 1498 A; 1665 C; 1485 G; 1405 T; 8 other;

Query Match 39.1%; Score 604.4; DB 18; Length 6061;
Best Local Similarity 91.4%; Pred. No. 3.7e-173;
Matches 638; Conservative 0; Mismatches 60; Indels 0; Gaps 0;
QY 595 TTACTGGCTGCTCCCTTATCGGGAAGCGGGCGATCATCAATGACGGCGCGCTG 654
Db 3948 ttatcatgtctggtccagatctcggcccatcgcgcgcgcatcgaatnnnnacatgtag 4007
QY 655 TAAAGTGTTAGCTTGAGAAAGATTCTGTCAGCCCGCGCTGCTGGCGTTTTCATTA 714
Db 4008 caaagggccagcaaaagccaggaacccgtaaaagggcgcgttgctgctgtttttcata 4067
QY 715 GGCTCGCCCGCCCTGACGAGCATCACAAAAATCGACGCTCAAGTCAGAGTGCGGAAACC 774
Db 4068 ggtccgcgccccctgacgagcatcacaaaaatcgacgctcaagtcagagggcgcaaac 4127
QY 775 CGACAGACTATAAAGATACAGCGGTTTCCCGCTGGAAGCTCCCTCGTGCCTCTCTG 834
Db 4128 cgacagactataaagataccagggctttccctcctggaagctcctcgtcgctcctcg 4187
QY 835 TTCGACCCCTGCGCTTACCGGATACCTGTCGCCCTTCTCCCTTCGGGAGCGTGCGCG 894
Db 4188 ttccgacctgcgcgttacccggatacctgtccgcctttcccttcggggaagcgtggcgc 4247
QY 895 TTTCTCAATGCTCAGCGCTGAGGTATCTCAGTTTCGGGTGAGGTCTGCTCCAAAGCTGG 954
Db 4248 ttctcaatgctcacgctgtaggtatctcagttcgggtgtaggtcgtcccaagctgg 4307
QY 955 GCTGTGTGACGAACCCCGCTTACGCCCGACCGCTCGCGCTTATCCGGTAACATATGTC 1014
Db 4308 gctgtgtgacgaaccccccttcagccgacccgtgcgcttatccgctaactatcgtc 4367
QY 1015 TTGAGTCCAAACCCGGTAAAGACACACTTATCGCCACTGCGCAGCAGCAGCTGCTAACAGGA 1074
Db 4368 ttgagtcacaacccggtaagacacgacttatcgcaactggcagcagccactggtaacagga 4427
QY 1075 TTAGCAGAGCAGGTATGATAGCGGTGCTACAGAGTTCTTGAAGTGGTGCCTTAACACTACG 1134
Db 4428 ttgacagagcgaggtatgtagggcgggtgtaggtccttgaagtggtggtcccaaccacg 4487
QY 1135 GCTACACTAGAAGGACACAGTATTTGGTATCTGCGCTCTGCTGAAGCCAGTTACCTTCGGAA 1194
Db 4488 gctacactagaagacagatttggtagtctgctcgtcgtgaagccagttacacttcgaa 4547
QY 1195 AAGAGTTGTTAGCTCTTGATCCGGCAACAAACACCGCTGCTGAGCGGTGTTTTTTG 1254
Db 4548 aagagttgtagctcttgatccgcgcaaacacacccgcgtgtagcgttggttttttg 4607
QY 1255 TTGCAAGCAGCAGATTACCGCGAGAAAAAAGGATCT 1292
Db 4608 ttgcaagcagcagattacgcgcagaaaaaaggatct 4645

RESULT 15
ID AAT90691 standard; DNA; 6312 BP.
XX
AC AAT90691;
XX
DT 05-JAN-1998 (first entry)
XX
DE Plasmid FBdelPMOSAF coding sequence.
XX
KW Packaging-deficient construct; viral gag-pol gene; packaging cell line;
KW moloney murine leukaemia virus; MOMLV; viral env gene; helper construct;
KW gene therapy; ss.
XX
OS Synthetic.
XX
PN WO9708330-A1.
XX

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: January 17, 2002, 11:54:39 ; Search time 221.34 Seconds
(without alignments)
1582.911 Million cell updates/sec

Title: US-09-242-202a-27
Perfect score: 1547
Sequence: 1 GGTACTGTCACCATGGCGC.....CTGCACAGCCTCTCCACA 1547

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 351203 seqs, 113238999 residues

Total number of hits satisfying chosen parameters: 702406

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued_Patents_NA:*
1: /cgn2.6/ptodata/2/ina/5A_COMB.seq:*
2: /cgn2.6/ptodata/2/ina/5B_COMB.seq:*
3: /cgn2.6/ptodata/2/ina/6A_COMB.seq:*
4: /cgn2.6/ptodata/2/ina/6B_COMB.seq:*
5: /cgn2.6/ptodata/2/ina/PTUS_COMB.seq:*
6: /cgn2.6/ptodata/2/ina/backfiles1.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	604.4	39.1	5865	4	US-09-011-745-8
2	604.4	39.1	6028	4	US-09-011-745-5
3	604.4	39.1	6061	4	US-09-011-745-6
4	604.4	39.1	6312	4	US-09-011-745-7
5	604.2	39.1	13254	1	US-08-276-852-156
6	604.2	39.1	13254	1	US-08-276-852-170
7	604.2	39.1	13254	1	US-08-899-575-156
8	604.2	39.1	13254	1	US-08-899-575-170
9	604.2	39.1	13254	1	US-08-899-575-156
10	604.2	39.1	13254	1	US-08-899-575-170
11	604.2	39.1	13254	5	PCT-US95-08743-156
12	604.2	39.1	13254	5	PCT-US95-08743-170
13	603	39.0	1305	1	US-08-594-469-9
14	603	39.0	1905	2	US-08-906-957-9
15	603	39.0	2927	2	US-08-941-647A-1
16	603	39.0	2939	1	US-08-119-512-2
17	603	39.0	2939	1	US-08-488-015B-2
18	603	39.0	2939	3	US-08-814-412-17
19	603	39.0	3003	6	5182260-18
20	603	39.0	3104	1	US-07-415-307A-1
21	603	39.0	3104	1	US-08-371-320-1
22	603	39.0	3130	4	US-09-038-141-1
23	603	39.0	3301	2	US-08-447-430A-42
24	603	39.0	3423	2	US-08-447-430A-40
25	603	39.0	3474	2	US-08-447-430A-41
26	603	39.0	3474	2	US-08-318-837-10
27	603	39.0	3585	1	US-08-362-670B-9

C 28	603	39.0	3585	3	US-08-333-576C-9	Sequence 9, Appli
C 29	603	39.0	3585	4	US-08-808-324-9	Sequence 9, Appli
C 30	603	39.0	3585	5	PCT-US94-14030A-9	Sequence 9, Appli
C 31	603	39.0	3623	2	US-07-989-847-13	Sequence 13, Appli
C 32	603	39.0	3623	4	US-08-469-411-13	Sequence 13, Appli
C 33	603	39.0	3632	1	US-07-745-382-13	Sequence 13, Appli
C 34	603	39.0	3632	1	US-08-115-680-13	Sequence 13, Appli
C 35	603	39.0	3632	1	US-07-941-372-3	Sequence 3, Appli
C 36	603	39.0	3632	1	US-08-165-301A-13	Sequence 13, Appli
C 37	603	39.0	3632	3	US-08-810-436-13	Sequence 13, Appli
C 38	603	39.0	3632	5	PCT-US93-08247-3	Sequence 3, Appli
C 39	603	39.0	3632	5	PCT-US94-14179-13	Sequence 13, Appli
C 40	603	39.0	3632	1	US-08-232-463-8	Sequence 8, Appli
C 41	603	39.0	3656	1	US-08-232-463-9	Sequence 9, Appli
C 42	603	39.0	3688	1	US-09-075-019-8	Sequence 8, Appli
C 43	603	39.0	3789	4	US-07-623-953-1	Sequence 1, Appli
C 44	603	39.0	3803	1	US-08-675-566-8	Sequence 8, Appli
C 45	603	39.0	3822	3		

ALIGNMENTS

RESULT 1
US-09-011-745-8
; Sequence 8, Application US/09011745
; Patent No. 6165715
; GENERAL INFORMATION:
; APPLICANT: Collins, Mary KL
; APPLICANT: Weiss, Robin A
; APPLICANT: Takeuchi, Yasuhiro
; APPLICANT: Cosset, Francois-Loic
; TITLE OF INVENTION: Expression systems
; FILE REFERENCE: 09/011,745
; CURRENT APPLICATION NUMBER: US/09/011,745
; EARLIER FILING DATE: 1998-06-22
; EARLIER APPLICATION NUMBER: PCT/GB96/02061
; EARLIER FILING DATE: 1996-08-23
; EARLIER APPLICATION NUMBER: GB9517263.1
; EARLIER FILING DATE: 1995-08-23
; NUMBER OF SEQ ID NOS: 29
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 8
; LENGTH: 5865
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: Portion of
; OTHER INFORMATION: Construct
; FEATURE:
; NAME/KEY: misc.feature
; LOCATION: (3611)
; OTHER INFORMATION: n is any nucleotide
; FEATURE:
; NAME/KEY: misc.feature
; LOCATION: (3612)
; OTHER INFORMATION: n is any nucleotide
; FEATURE:
; NAME/KEY: misc.feature
; LOCATION: (3613)
; OTHER INFORMATION: n is any nucleotide
; FEATURE:
; NAME/KEY: misc.feature
; LOCATION: (3614)
; OTHER INFORMATION: n is any nucleotide
; FEATURE:
; NAME/KEY: misc.feature
; LOCATION: (3799)
; OTHER INFORMATION: n is any nucleotide
; FEATURE:
; NAME/KEY: misc.feature
; LOCATION: (3800)
; OTHER INFORMATION: n is any nucleotide


```

; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (3801)
; OTHER INFORMATION: n is any nucleotide
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (3802)
; OTHER INFORMATION: n is any nucleotide
US-09-011-745-8

```

```

Query Match      39.1%; Score 604.4; DB 4; Length 5865;
Best Local Similarity 91.4%; Pred. No. 6.2e-169;
Matches 638; Conservative 0; Mismatches 60; Indels 0; Gaps 0;

```

```

QY 595 TTTACTGGCTGCTCCCTTATCGGAGGCGGCGCATCATATCAATGACGCGCGCTG 654
Db 3752 ttatcatgtctgcatcagatctggcccatcgccgcgcatcgtatnnnncatgtgag 3811
QY 655 TAAAGTGTACGTTGAGAAAGAAATTCCTGAGCCCGCGGTTGCTGGGCTTTTCCATA 714
Db 3812 caaagggccagcaaaagccaggaacgcgttaaagggcggtgtggtgttttccata 3871
QY 715 GGCTCCGCCCTCCCTGACGAGCATCAAAAATCGACGCTCAAGTCAGAGGTGCGGAACC 774
Db 3872 gctcgcgcgcctgagcgagcatcaaaaatcgagcgtcaagtcagagtggtgcaaac 3931
QY 775 CGACGAGCTATTAAGATACAGCGGCTTTCCCTGGAAGCTCCCTGCGCTCTCTCTG 834
Db 3932 cgacaggactataaagataccagcggtttcccccgtggaagctccctcggtctctc 3991
QY 835 TTCCGACCTGCGGCTTACCGATACCTGTCGCCCTTTCTCCCTGCGGAAGCGTGGCG 894
Db 3992 ttccgacctgacctgacgataacctgctccgctttctccctcggaagcggtggcg 4051
QY 895 TTTCTCAATGCTCACGCTGCTAGTATCTCAGTTCGGTGTAGCTGCTTCCCTCCAAGCT 954
Db 4052 ttctctcaatgctcacgctgtagtattctcagttcggtgtagtctgctccaaagctg 4111
QY 955 GCTGTGTGACGAACCCCGCTTACGCGGCGGCTTACCGGCTTATCCGCTAACTTCGTC 1014
Db 4112 gctgtgtgacgaaccccccttgcagccgacgctgagcttattccgtaactatgctc 4171
QY 1015 TTGAGTCCAAACCGGTAAACAGAGCTTATCGCCACTGCGCAGCAGCACTGTGTAAAC 1074
Db 4172 ttgagttccaaacccggttaacagagcttattcgccactgagcagcagcaggtgaa 4231
QY 1075 TTAGCAGAGGAGGTATGATGAGCGGTGTACAGAGTCTTGAAGTGTGGCTTAAGTAC 1134
Db 4232 ttgacagagaggtatgtaggggtgtctacagagttcttgagtggtggtggtggtg 4291
QY 1135 GTTACTAGAGGAGGAGTATTTGTTATCTGCGCTGTGCTGAAGCAGTTACTTTCGGA 1194
Db 4292 gctactactagaagagcagatttggatctgctgctgctgagcaggttactcttcgaa 4351
QY 1195 AAAGAGTGTGAGCTTTCATCCGCAACAAACACACCGCTGCTGAGCGTGTGTTTTT 1254
Db 4352 aaagagttggtgagctgtgagtcgagcaaacacacccgctggtgagcgtgtgtttt 4411
QY 1255 TTTGCAAGCAGCAGATTACGCGCAGAAAAGAGATCT 1292
Db 4412 ttgcaagcagcagattacgcgcagaaaaaaggatct 4449

```

```

RESULT 2
US-09-011-745-5
; Sequence 5, Application US/09011745
; Patent No. 6165715
; GENERAL INFORMATION:
; APPLICANT: Collins, Mary KL
; APPLICANT: Weiss, Robin A
; APPLICANT: Takeuchi, Yasuhiro
; APPLICANT: Cosset, Francois-Loic

```

```

; TITLE OF INVENTION: Expression systems
; FILE REFERENCE: 09/011,745
; CURRENT APPLICATION NUMBER: US/09/011,745
; EARLIER FILING DATE: 1998-06-22
; EARLIER APPLICATION NUMBER: PCT/GB96/02061
; EARLIER FILING DATE: 1996-08-23
; EARLIER APPLICATION NUMBER: GB9517263.1
; EARLIER FILING DATE: 1995-08-23
; NUMBER OF SEQ ID NOS: 29
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 5
; LENGTH: 6028
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: Portion of
; OTHER INFORMATION: construct
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (3774)
; OTHER INFORMATION: n is any nucleotide
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (3775)
; OTHER INFORMATION: n is any nucleotide
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (3776)
; OTHER INFORMATION: n is any nucleotide
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (3777)
; OTHER INFORMATION: n is any nucleotide
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (3962)
; OTHER INFORMATION: n is any nucleotide
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (3963)
; OTHER INFORMATION: n is any nucleotide
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (3964)
; OTHER INFORMATION: n is any nucleotide
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (3965)
; OTHER INFORMATION: n is any nucleotide
US-09-011-745-5

```

```

Query Match      39.1%; Score 604.4; DB 4; Length 6028;
Best Local Similarity 91.4%; Pred. No. 6.3e-169;
Matches 638; Conservative 0; Mismatches 60; Indels 0; Gaps 0;
QY 595 TTTACTGGCTGCTCCCTTATCGGAGGCGGCGCATCATATCAATGACGCGCGCTG 654
Db 3915 ttatcatgtctgcatcagatctggcccatcgccgcgcatcgtatnnnncatgtgag 3974
QY 655 TAAAGTGTACGTTGAGAAAGAAATTCCTGAGCCCGCGGTTGCTGGGCTTTTCCATA 714
Db 3975 caaagggccagcaaaagccaggaacgcgttaaagggcggtgtggtggttttccata 4034
QY 715 GGCTCCGCCCTCCCTGACGAGCATCAAAAATCGACGCTCAAGTCAGAGGTGCGGAACC 774
Db 4035 ggtcctccgcgcctcgagcagcatcaaaaatcgacgtcaagtcagagtggtgcaaac 4094
QY 775 CGACGAGCTATTAAGATACAGCGGCTTTCCCTGGAAGCTCCCTGCGCTCTCTCTG 834
Db 4095 cgacagagactataaagataccagcggtttcccccgtggaagctccctcgctcctg 4154
QY 835 TTCCGACCTGCGGCTTACCGGATACCTGTGCGGCTTTCTCCCTTCGGAAGCGTGGCG 894

```



```

; OTHER INFORMATION: n is any nucleotide
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (3995)
; OTHER INFORMATION: n is any nucleotide
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (3996)
; OTHER INFORMATION: n is any nucleotide
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (3997)
; OTHER INFORMATION: n is any nucleotide
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (3998)
; OTHER INFORMATION: n is any nucleotide
US-09-011-745-6

Query Match      39.1%; Score 604.4; DB 4; Length 6061;
Best Local Similarity 91.4%; Pred. No. 6.3e-169;
Matches 638; Conservative 0; Mismatches 60; Indels 0; Gaps 0;

Qy 595 TTTRACTGGCTGCTCCCTTATCGGGAGAGCGGGGGGCATCATATCAAAATGACGCGCGCTG 654
Db 3948 ttatcatgtctggatccagatctggcccatcgccgcgagatcgatnnnnacatgtgag 4007
Qy 655 TAAAGTGTACGTTGAGAAAAGAAATTCCTGCGAGCCCGCGGCTGCTGGCGTTTTCCTATA 714
Db 4008 caaaagccagcaaaagccaggaacccgtataaaagccggtgtgctgtgctgttttcata 4067
Qy 715 GGCTCCGCGCCCTGAGCAGCATCACAAAATCGACCTCAAGTCAGAGTGGCGGAACC 774
Db 4068 ggtctccgccccctgagcgatcacaaaaatcgacgtcaagtcagaggtggcgaaacc 4127
Qy 775 CGCAGGACTATAAAGATACACAGCGCTTCCCTCGAAGCTCCCTGCTGGCTCTCTCTG 834
Db 4128 cgacgactataaagataccagcggttcccccctggaagctccctcgtgctctcctg 4187
Qy 835 TTCGACCCCTGCGCTTACCGGATACCTGCGCGCTTTCCTCCCTTCGGGAAGCGTGGCG 894
Db 4188 ttcgagccctgcccctaccggatcacgtgtccgccttctccctcctgggaagcggtgc 4247
Qy 895 TTTCTCAATGCTACGCTGATAGGTATCTCAGTTCGCTGCTAGGTGCTTCCGCTTCCAGCTGG 954
Db 4248 ttctcaatgctcacgctgtaggatctcagttcgttctcgttccctcccaagctgg 4307
Qy 955 GCTGTGTCAGCAACCCCTTCAGCCCGGCTGCGCTTATCGCGGTAACCTATCGTC 1014
Db 4308 gctgtgtgacgaaccccccttcagccccgcgctgagccttaccgttaactatcgtc 4367
Qy 1015 TTGAGTCCAAACCCGGTAAGACACACTTATCGCCACCTGGCAGCAGCCACTGGTAACAGGA 1074
Db 4368 ttgagtcacaacccggtaagacacgacttatcgccacttggcagcagccactggttaacagga 4427
Qy 1075 TTACGACGCGAGCTATGTAGCGGCTGCTACAGAGTTCCTGAAGTGGTGGCTAACTACG 1134
Db 4428 ttacgagcgaggtatgtaggcggtgctacagagttcttgaagtgggtggcctaactacg 4487
Qy 1135 GCTACACTAGAAAGGACAGTATTTGGTATCTCGCTGCTGCTGAAGCCAGTTACCTTCGGAA 1194
Db 4488 gctacactagaaggaagatatttgatctgctgtcgtctgctgaagccagttacccttcgaa 4547
Qy 1195 AAAGAGTTGGTAGCTCTTGATCCGGCAAAACACCGCTGGTAGCGGTGGTTTTTTT 1254
Db 4548 aaagagttggtagctcttgatcccggaacaaacacccgctggtagcggtgtgtttttg 4607
Qy 1255 TTTGCAAGCAGCAGATTACGCGCGAGAAAAAGGATCT 1292
Db 4608 ttgcaagcagcagattacgcgcagaaaaaaggatct 4645

```

RESULT 3

```

US-09-011-745-6
; Sequence 6, Application US/09011745
; Patent No. 6165715
; GENERAL INFORMATION:
; APPLICANT: Collins, Mary KL
; APPLICANT: Weiss, Robin A
; APPLICANT: Takeuchi, Yasuhiro
; APPLICANT: Cosset, Francois-Loic
; TITLE OF INVENTION: Expression systems
; FILE REFERENCE: 09/011,745
; CURRENT APPLICATION NUMBER: US/09/011,745
; CURRENT FILING DATE: 1998-06-22
; EARLIER APPLICATION NUMBER: PCT/GB96/02061
; EARLIER FILING DATE: 1996-08-23
; EARLIER APPLICATION NUMBER: GB9517263.1
; EARLIER FILING DATE: 1995-08-23
; NUMBER OF SEQ ID NOS: 29
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 6
; LENGTH: 6061
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: Portion of
; OTHER INFORMATION: construct
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (3807)
; OTHER INFORMATION: n is any nucleotide
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (3808)
; OTHER INFORMATION: n is any nucleotide
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (3809)
; OTHER INFORMATION: n is any nucleotide
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (3810)

```

```

RESULT 4
US-09-011-745-7
; Sequence 7, Application US/09011745
; Patent No. 6165715
; GENERAL INFORMATION:
; APPLICANT: Collins, Mary KL
; APPLICANT: Weiss, Robin A
; APPLICANT: Takeuchi, Yasuhiro
; APPLICANT: Cosset, Francois-Loic
; TITLE OF INVENTION: Expression systems
; FILE REFERENCE: 09/011,745
; CURRENT APPLICATION NUMBER: US/09/011,745
; EARLIER FILING DATE: 1998-06-22
; EARLIER APPLICATION NUMBER: PCT/GB96/02061
; EARLIER FILING DATE: 1996-08-23
; EARLIER APPLICATION NUMBER: GB9517263.1
; EARLIER FILING DATE: 1995-08-23
; NUMBER OF SEQ ID NOS: 29
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 7
; LENGTH: 6312
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: Portion of
; OTHER INFORMATION: construct
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (4058)
; OTHER INFORMATION: n is any nucleotide
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (4059)
; OTHER INFORMATION: n is any nucleotide
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (4060)
; OTHER INFORMATION: n is any nucleotide
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (4061)
; OTHER INFORMATION: n is any nucleotide
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (4246)
; OTHER INFORMATION: n is any nucleotide
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (4247)
; OTHER INFORMATION: n is any nucleotide
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (4248)
; OTHER INFORMATION: n is any nucleotide
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (4249)
; OTHER INFORMATION: n is any nucleotide
US-09-011-745-7

Query Match 39.1%; Score 604.4; DB 4; Length 6312;
Best Local Similarity 91.4%; Pred. NO. 6.4e-169;
Matches 638; Conservative 0; Mismatches 60; Indels 0; Gaps 0;

QY 595 TTACTGGCTGCTCCTTATCGGAAGCGGCGCATCATCAAAATGACGCGCGCTG 654
|| | ||| || | ||| || | ||| || | ||| || | ||| || | ||| || |
Db 4199 ttatgtctgtgattccagatctggcccatcgccgcgagatctatnnnacaatgtga 4258
|| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |

QY 655 TAAAGTGTTCAGTTGAGAAAGAAATCTCGACGCCCGCGCTTGTGCGGTTTTCATA 714
|| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 4259 caaaagccagcaaaagccaggaaaccgtaaaagccgctgtgctggcggttttccata 4318
|| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |

```

```

QY 715 GCTCCGCCCCCTGACGAGATCACAAAAATCGACGCTCAAGTCAGAGGTGGCGAACCC 774
|| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 4319 ggctccgccccctgaogagcatcacaaaaatcgacgctcaagtcagaggtggcgaaacc 4378
|| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 775 CGACAGGACTATAAGATACAGAGCGGTTTCCCTCGGAAGCTCCCTCGTGGCTCTCCTG 834
|| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 4379 cgacaggactataaagataaccaggcggtttcccccccggaagctccctctgctcctg 4438
|| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 835 TTCCGACCTGCGGCTTACCGGATACCTGTCCGCCCTTTCTCCCTTCGGAAGCGTGGCGC 894
|| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 4439 ttccgaacctgccgtaccggataacctgtccgctttctcccttcgggaagcgtggcgc 4498
|| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 895 TTCTCAATGCTACGCTGCTAGGTATCTCAGTTCCGTTAGGTGCTGCTCCCAAGCTGG 954
|| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 4499 ttctcaatgctcaogctgtaggtatcagttcggttaggtcggtcccaagcgtggcgc 4558
|| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 955 GCTGTGTCGACGAACCCCGCTTCAGCGCGACCGCTGCGCCCTTATCCGCTAACTATCCTC 1014
|| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 4559 gctgtgtgcaagaaacccccggttcagcccgaccgctgctgcttaccggttaactatcgtc 4618
|| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 1015 TTGAGTCCAAACCCGGTAAAGACACGACTTATCCGCTGCGACGACGCTGCTTAACAGA 1074
|| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 4619 ttgagtcgaacccggtaagacacgacttatcgccactggcagcgcactggtaacagga 4678
|| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 1075 TTAGCAGACGAGGTATGTAGCGGTGCTACAGAGTCTTGAAGTGGTGGCTTAACAG 1134
|| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 4679 ttgacagagcgaggtatgtagcggtgtacagagttcttggaagtgggtggttaactacg 4738
|| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 1135 GCTACACTAGAAGGACAGTATTGGTATCTGCGCTGCTGCTGAAGCCAGTACCTTCGGA 1194
|| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 4739 gctacactagaagagcagatttgggtatgtatgctgctgctgctgaagccagttaccttcgga 4798
|| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 1195 AAAGAGTTGTAGTCTTGTATCGGCAAAACAAACACCGCTGCTAGCGGTGTTTTTTT 1254
|| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 4799 aaagagttgtagctctgtatcgcgcaaaacacccgctgtagcggtggtttttt 4858
|| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 1255 TTGCAAGCAGCAGATTACGCGCAGAAAAAAGGATCT 1292
|| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 4859 ttgcaagcagcagattacgcgcaaaaaaaggatct 4896
|| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |

```

```

RESULT 5
US-08-276-852-156
; Sequence 156, Application US/08276852
; Patent No. 5652138
; GENERAL INFORMATION:
; APPLICANT: Burton, Dennis R
; APPLICANT: Barbas, Carlos F
; APPLICANT: Lerner, Richard A
; TITLE OF INVENTION: HUMAN NEUTRALIZING MONOCLONAL ANTIBODIES
; TITLE OF INVENTION: TO HUMAN IMMUNODEFICIENCY VIRUS
; NUMBER OF SEQUENCES: 170
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: The Scripps Research Institute, Office of
; STREET: 10666 No. 5652138th Torrey Pines Road, Suite 220,
; CITY: La Jolla
; STATE: CA
; COUNTRY: USA
; ZIP: 92037
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/276,852
; FILING DATE: 18-JUL-1994
; CLASSIFICATION: 514
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/178,302
; FILING DATE: 30-SEP-1993

```

```

; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/954,148
; FILING DATE: 30-SEP-1992
; ATTORNEY/AGENT INFORMATION:
; NAME: Fitting, Thomas
; REGISTRATION NUMBER: 34,163
; REFERENCE/DOCKET NUMBER: SCRI452P
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 619-554-2937
; TELEFAX: 619-554-6312
; INFORMATION FOR SEQ ID NO: 156:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 13254 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: circular
; MOLECULE TYPE: DNA (genomic)
; US-08-276-852-156

Query Match          39.1%; Score 604.2; DB 1; Length 13254;
Best Local Similarity 97.9%; Pred. No. 1e-168;
Matches 612; Conservative 0; Mismatches 13; Indels 0; Gaps 0;

QY 668 TGAGAAAGAAATTCCTGCAGCCCGCGCTGTGCGCTTTTTCATAGGCTCCGCCCC 727
DB 5976 TAAGGGAGAGCGTGCACCTCGGGCGGTTGCTGGCGTTTTTCCATAGGCTCCGCCCC 6035

QY 728 TGAGGACATCACAAAAATCGAGCTCAAGTCAGAGTGGCGAAACCCGACAGACTATA 787
DB 6036 TGAGGACATCACAAAAATCGAGCTCAAGTCAGAGTGGCGAAACCCGACAGACTATA 6095

QY 788 AAGATACAGCGGTTTCCCTCGGAAGCTCCCTGCTGCGCTCTCTCTGTCGACCCGTC 847
DB 6096 AAGATACAGCGGTTTCCCTCGGAAGCTCCCTGCTGCGCTCTCTCTGTCGACCCGTC 6155

QY 848 GCTTACCGGATACCTGTCGCGCTTTCTCCCTTCGGGAAGCGTGGCGCTTTCTCAATGCTC 907
DB 6156 GCTTACCGGATACCTGTCGCGCTTTCTCCCTTCGGGAAGCGTGGCGCTTTCTCAATGCTC 6215

QY 908 ACGCTGTAGGTATCTCAGTTCGGTGTAGTTCGCTCCAAAGCTGGCGTGTGTGCACGA 967
DB 6216 ACGCTGTAGGTATCTCAGTTCGGTGTAGTTCGCTCCAAAGCTGGCGTGTGTGCACGA 6275

QY 968 ACCCCCGTTTCAGCCGACCGCTCGCTTATCCGGTAACTATGCTCTGAGTCCAAACC 1027
DB 6276 ACCCCCGTTTCAGCCGACCGCTCGCTTATCCGGTAACTATGCTCTGAGTCCAAACC 6335

QY 1028 GGTAAAGACAGACTTATCGCCACTGGCAGCAGCCACTGGTAACAGGATTAGCAGAGCGAG 1087
DB 6336 GGTAAAGACAGACTTATCGCCACTGGCAGCAGCCACTGGTAACAGGATTAGCAGAGCGAG 6395

QY 1088 GTATGTAGCGGCTGCTACAGAGTTCTTGAAGTGTGGCTTAACCTACGGCTACACTAGAAG 1147
DB 6396 GTATGTAGCGGCTGCTACAGAGTTCTTGAAGTGTGGCTTAACCTACGGCTACACTAGAAG 6455

QY 1148 GACAGTATTTGGTATCGCTCGCTGCTGAAACCCAGTTACCTTCGGAAGAGTTGGTAG 1207
DB 6456 GACAGTATTTGGTATCGCTCGCTGCTGAAACCCAGTTACCTTCGGAAGAGTTGGTAG 6515

QY 1208 CTCCTGTATCGGCAAAACCAACCGCTGTGTAGGGTGGTTTTTTTGTTCGAAGACGACGA 1267
DB 6516 CTCCTGTATCGGCAAAACCAACCGCTGTGTAGGGTGGTTTTTTTGTTCGAAGACGACGA 6575

QY 1268 GATTACCGCGAGAAAAAAGATCT 1292
DB 6576 GATTACCGCGAGAAAAAAGATCT 6600

RESULT 6
US-08-276-852-170/c
; Sequence 170, Application US/08276852
; Patent No. 5652138

```

```

; GENERAL INFORMATION:
; APPLICANT: Dennis R
; APPLICANT: Burton, Dennis R
; APPLICANT: Barbas, Carlos F
; APPLICANT: Lerner, Richard A
; TITLE OF INVENTION: HUMAN NEUTRALIZING MONOCLONAL ANTIBODIES
; TITLE OF INVENTION: TO HUMAN IMMUNODEFICIENCY VIRUS
; NUMBER OF SEQUENCES: 170
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: The Scripps Research Institute, Office of
; ADDRESSEE: Patent Counsel
; STREET: 10666 No. 5652138th Torrey Pines Road, Suite 220,
; STREET: Mail Drop TPC8
; CITY: La Jolla
; STATE: CA
; COUNTRY: USA
; ZIP: 92037
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA: US/08/276,852
; FILING DATE: 18-JUL-1994
; CLASSIFICATION: 514
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/178,302
; FILING DATE: 30-SEP-1993
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/954,148
; FILING DATE: 30-SEP-1992
; ATTORNEY/AGENT INFORMATION:
; NAME: Fitting, Thomas
; REGISTRATION NUMBER: 34,163
; REFERENCE/DOCKET NUMBER: SCRI452P
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 619-554-2937
; TELEFAX: 619-554-6312
; INFORMATION FOR SEQ ID NO: 170:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 13254 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: circular
; MOLECULE TYPE: DNA (genomic)
; US-08-276-852-170

```

```

Query Match          39.1%; Score 604.2; DB 1; Length 13254;
Best Local Similarity 97.9%; Pred. No. 1e-168;
Matches 612; Conservative 0; Mismatches 13; Indels 0; Gaps 0;

QY 668 TGAGAAAGAAATTCCTGCAGCCCGCGCTGTGCGCTTTTTCATAGGCTCCGCCCC 727
DB 7279 TAAGGGAGAGCGTGCACCTCGGGCGGTTGCTGGCGTTTTTCCATAGGCTCCGCCCC 7220

QY 728 TGAGGACATCACAAAAATCGAGCTCAAGTCAGAGTGGCGAAACCCGACAGACTATA 787
DB 7219 TGAGGACATCACAAAAATCGAGCTCAAGTCAGAGTGGCGAAACCCGACAGACTATA 7160

QY 788 AAGATACAGCGGTTTCCCTCGGAAGCTCCCTGCTGCGCTCTCTGTCGACCCGTC 847
DB 7159 AAGATACAGCGGTTTCCCTCGGAAGCTCCCTGCTGCGCTCTCTGTCGACCCGTC 7100

QY 848 GCTTACCGGATACCTGTCGCGCTTTCTCCCTTCGGGAAGCGTGGCGCTTTCTCAATGCTC 907
DB 7099 GCTTACCGGATACCTGTCGCGCTTTCTCCCTTCGGGAAGCGTGGCGCTTTCTCAATGCTC 7040

QY 908 ACGCTGTAGGTATCTCAGTTCGGTGTAGTTCGCTCCAAAGCTGGCGTGTGTGCACGA 967
DB 7039 ACGCTGTAGGTATCTCAGTTCGGTGTAGTTCGCTCCAAAGCTGGCGTGTGTGCACGA 6980

QY 968 ACCCCCGTTTCAGCCGACCGCTCGCTTATCCGGTAACTATGCTCTGAGTCCAAACC 1027

```



```

; ZIP: 92037
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/899,575
; FILING DATE: 24-JUL-1997
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/276,852
; FILING DATE: 18-JUL-1994
; APPLICATION NUMBER: US 08/178,302
; FILING DATE: 30-SEP-1993
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/954,148
; FILING DATE: 30-SEP-1992
; ATTORNEY/AGENT INFORMATION:
; NAME: Fitting, Thomas
; REGISTRATION NUMBER: 34,163
; REFERENCE/DOCKET NUMBER: SCR1452P
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 619-554-2937
; TELEFAX: 619-554-6312
; INFORMATION FOR SEQ ID NO: 170:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 13254 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: circular
; MOLECULE TYPE: DNA (genomic)
; US-08-899-575-170

Query Match          39.1%; Score 604.2; DB 1; Length 13254;
Best Local Similarity 97.9%; Pred. No. le-168;
Matches 612; Conservative 0; Mismatches 13; Indels 0; Gaps 0;

QY   668 TGAGAAAGAATTCTCGCAGCCGCCGGTGTGCTGGCGTTTTTCATAGGCTCCGCCCCCC 727
DB   7279 TAAGGAGAGCGTCGACCTCGGGCGGGTGCTGGCGTTTTTCATAGGCTCCGCCCCCC 7220

QY   728 TGACGAGCATCACAAAAATCGACGCTCAAGTTCAGAGGTGGCGAAACCCGACGAGACTATA 787
DB   7219 TGACGAGCATCACAAAATCGACGCTCAAGTCAGAGGTGGCGAAACCCGACGAGACTATA 7160

QY   788 AAGATACAGCGGTTCCCGCTTGAAGCTCCCTCGTGGCGTCTCCTGTTCGAGCCCTGCC 847
DB   7159 AAGATACAGCGGTTCCCGCTTGAAGCTCCCTCGTGGCGTCTCCTGTTCGAGCCCTGCC 7100

QY   848 GCATTACCGGTACTCTGCCTTCCTCCCTTCGGGAGCGTGGCGCTTTCTCAATGCTC 907
DB   7099 GCATTACCGGTACTCTGCCTTCCTCCCTTCGGGAGCGTGGCGCTTTCTCAATGCTC 7040

QY   908 ACGTGTAGGTATCTCAGTTCGGTGTAGTTCGCTCCAAGCTGGGCTGTGTGCACA 967
DB   7039 ACGTGTAGGTATCTCAGTTCGGTGTAGTTCGCTCCAAGCTGGGCTGTGTGCACA 6980

QY   968 ACCCCCCTTAGCCCCGACCGCTCGGCTTAACGGGTAACTATCGTCTTGAGTCCAACCC 1027
DB   6979 ACCCCCCTTAGCCCCGACCGCTCGGCTTAACGGGTAACTATCGTCTTGAGTCCAACCC 6920

QY   1028 GGTAAGACAGCACTTATCGCCACTGCGAGCAGCCACTGGTAACAGGATTACGAGCGGAG 1087
DB   6919 GGTAAGACAGCACTTATCGCCACTGCGAGCAGCCACTGGTAACAGGATTACGAGCGGAG 6860

QY   1088 GTATGTAGCGGGTGTACAGAGTTCTTTGAAGTGTGGCCTAACTACGGCTACACTAGAAG 1147
DB   6859 GTATGTAGCGGGTGTACAGAGTTCTTTGAAGTGTGGCCTAACTACGGCTACACTAGAAG 6800

QY   1148 GACAGTATTTGGTATCTGCGCTGTGCTGAAGCCAGTTACCTTCGGA AAAAAGATTGGTGA 1207

```

QY 728 TGACGAGCATCAAAAATCGAGCTCAAGTCAGAGTGGCGAAACCGACAGGACTATA 787
Db 6036 TGACGAGCATCAAAAATCGAGCTCAAGTCAGAGTGGCGAAACCGACAGGACTATA 6095
QY 788 AAGATACCGAGCTTTCCCGCTGGAAGCTCCCTGCGGCTCTCTCTGTTCCGACCCCTGCC 847
Db 6096 AAGATACCGAGCTTTCCCGCTGGAAGCTCCCTGCGGCTCTCTCTGTTCCGACCCCTGCC 6155
QY 848 GCTTACCGGATACCTGTCGCGCTTTCCCTTCGGAAGGCTGGCGCTTTCTCAATGCTC 907
Db 6156 GCTTACCGGATACCTGTCGCGCTTTCCCTTCGGAAGGCTGGCGCTTTCTCAATGCTC 6215
QY 908 ACCTGTAGTATCTCAGTTCCGCTGAGTGTCTGCTCCAAAGCTGGGCTGTGTGCAACGA 967
Db 6216 ACCTGTAGTATCTCAGTTCCGCTGAGTGTCTGCTCCAAAGCTGGGCTGTGTGCAACGA 6275
QY 968 ACCCCCGTTACCGCGAGCGCTGCGCTTATCCGCTAACTATCTGCTTCAGTCCAAACC 1027
Db 6276 ACCCCCGTTACCGCGAGCGCTGCGCTTATCCGCTAACTATCTGCTTCAGTCCAAACC 6335
QY 1028 GGTAAAGACGACTTATCGCCACTGGCAGCAGCCACTGGTAAACAGGATTAGCAGAGCGAG 1087
Db 6336 GGTAAAGACGACTTATCGCCACTGGCAGCAGCCACTGGTAAACAGGATTAGCAGAGCGAG 6395
QY 1088 GTATGTAGCGGTGTACAGAGTCTTGAAGTGGTGGCTTAACCTACGGCTACACTAGAAG 1147
Db 6396 GTATGTAGCGGTGTACAGAGTCTTGAAGTGGTGGCTTAACCTACGGCTACACTAGAAG 6455
QY 1148 GACAGTATTGGTATCTGCGCTCTGCTGAAGCCAGTACCTTCGGAAGAGTGTGGTAG 1207
Db 6456 GACAGTATTGGTATCTGCGCTCTGCTGAAGCCAGTACCTTCGGAAGAGTGTGGTAG 6515
QY 1208 CTCTTGATCGGCAAAACAAACCCAGCTGGTAGCGGTGGTTTTTTTGTTCGAAGCAGCA 1267
Db 6516 CTCTTGATCGGCAAAACAAACCCAGCTGGTAGCGGTGGTTTTTTTGTTCGAAGCAGCA 6575
QY 1268 GATTACGCGCAGAAAAAGGATCT 1292
Db 6576 GATTACGCGCAGAAAAAGGATCT 6600

RESULT 10

US-08-899-575-170/C
; Sequence 170, Application US/08899575
; Patent No. 5804440
; GENERAL INFORMATION:
; APPLICANT: Burton, Dennis R
; APPLICANT: Barbas, Carlos F
; APPLICANT: Lerner, Richard A
; TITLE OF INVENTION: HUMAN NEUTRALIZING MONOCLONAL ANTIBODIES
; NUMBER OF SEQUENCES: 170
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: The Scripps Research Institute, Office of
; ADDRESSER: Patent Counsel
; STREET: 10566 No. 580440th Torrey Pines Road, Suite 220,
; STREET: Mail Drop TPC8
; CITY: La Jolla
; STATE: CA
; COUNTRY: USA
; ZIP: 92037
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/899,575
; FILING DATE: 24-JUL-1997
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/276,852
; FILING DATE: 18-JUL-1994

; APPLICATION NUMBER: US 08/178,302
; FILING DATE: 30-SEP-1993
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/954,148
; FILING DATE: 30-SEP-1992
; ATTORNEY/AGENT INFORMATION:
; NAME: Fitting, Thomas
; REGISTRATION NUMBER: 34,163
; REFERENCE/DOCKET NUMBER: SCRI452P
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 619-554-2937
; TELEFAX: 619-554-6312
; INFORMATION FOR SEQ ID NO: 170:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 13254 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: circular
; MOLECULE TYPE: DNA (genomic)
US-08-899-575-170

Query Match 39.1%; Score 604.2; DB 1; Length 13254;
Best Local Similarity 97.9%; Pred. No. 1e-168;
Matches 612; Conservative 0; Mismatches 13; Indels 0; Gaps 0;

QY 668 TGAGAAAGAAATTCCTGCAGCCGCGCTTGTGCGGTTTTTCCATAGGCTCGGCCCC 727
Db 7279 TAAGGAGAGCGTCGACCTCGGGCGCGTGTGCGGTTTTTCCATAGGCTCGGCCCC 7220
QY 728 TGACGAGCATCAAAAATCGAGCTCAAGTCAGAGTGGCGAAACCGACAGGACTATA 787
Db 7219 TGACGAGCATCAAAAATCGAGCTCAAGTCAGAGTGGCGAAACCGACAGGACTATA 7160
QY 788 AAGATACCGAGGCTTCCCGCTGGAGCTCCCTCGCGCTCTCTCTGTTCCGACCCCTGCC 847
Db 7159 AAGATACCGAGGCTTCCCGCTGGAGCTCCCTCGCGCTCTCTCTGTTCCGACCCCTGCC 7100
QY 848 GCTTACCGGATACCTGTCGCGCTTTCTCCCTTCGGGAAGCGTGGCGCTTCTCAATGCTC 907
Db 7099 GCTTACCGGATACCTGTCGCGCTTTCTCCCTTCGGGAAGCGTGGCGCTTCTCAATGCTC 7040
QY 908 ACCTGTAGTATCTCAGTTCCGCTGAGTGTCTGCTCCAAAGCTGGGCTGTGTGCAACGA 967
Db 7039 ACCTGTAGTATCTCAGTTCCGCTGAGTGTCTGCTCCAAAGCTGGGCTGTGTGCAACGA 6980
QY 968 ACCCCCGTTACCGCGAGCGCTGGCGCTTATCCGCTAACTATCTGCTTGAAGTCCAAACC 1027
Db 6979 ACCCCCGTTACCGCGAGCGCTGGCGCTTATCCGCTAACTATCTGCTTGAAGTCCAAACC 6920
QY 1028 GGTAAAGACGACTTATCGCCACTGGCAGCAGCCACTGGTAAACAGGATTAGCAGAGCGAG 1087
Db 6919 GGTAAAGACGACTTATCGCCACTGGCAGCAGCCACTGGTAAACAGGATTAGCAGAGCGAG 6860
QY 1088 GTATGTAGCGGTGTACAGAGTCTTGAAGTGGTGGCTTAACCTACGGCTACACTAGAAG 1147
Db 6859 GTATGTAGCGGTGTACAGAGTCTTGAAGTGGTGGCTTAACCTACGGCTACACTAGAAG 6800
QY 1148 GACAGTATTGGTATCTGCGCTCTGCTGAAGCCAGTACCTTCGGAAGAGATTGGTAG 1207
Db 6799 GACAGTATTGGTATCTGCGCTCTGCTGAAGCCAGTACCTTCGGAAGAGATTGGTAG 6740
QY 1208 CTCTTGATCGGCAAAACAAACCCAGCTGGTAGCGGTGGTTTTTTTGTTCGAAGCAGCA 1267
Db 6739 CTCTTGATCGGCAAAACAAACCCAGCTGGTAGCGGTGGTTTTTTTGTTCGAAGCAGCA 6680
QY 1268 GATTACGCGCAGAAAAAGGATCT 1292
Db 6679 GATTACGCGCAGAAAAAGGATCT 6655

RESULT 11
PCT-US95-08743-156

; Sequence 156, Application PC/TUS9508743
; GENERAL INFORMATION:
; APPLICANT:
; TITLE OF INVENTION: HUMAN NEUTRALIZING MONOCLONAL ANTIBODIES
; TITLE OF INVENTION: TO HUMAN IMMUNODEFICIENCY VIRUS
; NUMBER OF SEQUENCES: 170
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25 (EPO)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: PCT/US95/08743
; FILING DATE: 11-JUL-1995
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/276,852
; FILING DATE: 18-JUL-1994
; INFORMATION FOR SEQ ID NO: 156:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 13254 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: circular
; MOLECULE TYPE: DNA (genomic)
; PCT-US95-08743-156

Query Match 39.1%; Score 604.2; DB 5; Length 13254;
Best Local Similarity 97.9%; Pred. No. 1e-168;
Matches 612; Conservative 0; Mismatches 13; Indels 0; Gaps 0;

QY	668	TGAGAAAGATTCTCGCAGCGCGCGTGGTGGCTTTTCCATAGGCTCGCCGCC	727
Db	5976	TAGGGAGAGCGTCGACCTCGGCGCGGTGCTGGCGTTTTTCCATAGGCTCGCCGCC	6035
QY	728	TGACGAGCATCACAAAATCGAGCTCAAGTCAGAGGTGGCGAAACCCGACAGACTATA	787
Db	6036	TGACGAGCATCACAAAATCGAGCTCAAGTCAGAGGTGGCGAAACCCGACAGACTATA	6095
QY	788	AAGATACCGAGCGTTCCTCCCTCGGAAGCTCCCTCGCGCTCTCTGTCGACCTGCC	847
Db	6096	AAGATACCGAGCGTTCCTCCCTCGGAAGCTCCCTCGCGCTCTCTGTCGACCTGCC	6155
QY	848	GCTTACCGGATACCTGTCCGCTTCTCCCTCGGAAGCTGGCGCTTCTCAATGCTC	907
Db	6156	GCTTACCGGATACCTGTCCGCTTCTCCCTCGGAAGCTGGCGCTTCTCAATGCTC	6215
QY	908	ACGCTGTAGTATCTCAGTTCGGTGTAGTGTGCTTCCAAAGCTGGGCTGTGTGCACGA	967
Db	6216	ACGCTGTAGTATCTCAGTTCGGTGTAGTGTGCTTCCAAAGCTGGGCTGTGTGCACGA	6275
QY	968	ACCCCGGTTACGCGGACCGCTGGCGCTTATCCGCTAAGTATCGTCTTCAATGCTC	1027
Db	6276	ACCCCGGTTACGCGGACCGCTGGCGCTTATCCGCTAAGTATCGTCTTCAATGCTC	6335
QY	1028	GGTAAGACAGCACTATTCGCACTGGCAGCAGCCACTGGTAACAGGATTAGCAGAGCGAG	1087
Db	6336	GGTAAGACAGCACTATTCGCACTGGCAGCAGCCACTGGTAACAGGATTAGCAGAGCGAG	6395
QY	1088	GTATGTAGGCGGTGTACAGAGTTCCTTGAAGTGGTGGCTTAACCTACGGCTACACTAGAAG	1147
Db	6396	GTATGTAGGCGGTGTACAGAGTTCCTTGAAGTGGTGGCTTAACCTACGGCTACACTAGAAG	6455
QY	1148	GACAGTATTTGGTATCTGCTGAGCCAGTACCTTCCGAAAGAGATTGGTAG	1207
Db	6456	GACAGTATTTGGTATCTGCTGAGCCAGTACCTTCCGAAAGAGATTGGTAG	6515
QY	1208	CTCTTGATCCGCAACAAACACCGCTGCTAGCGGTGGTGTCTTTTGTTCGAAGCAGCA	1267
Db	6516	CTCTTGATCCGCAACAAACACCGCTGCTAGCGGTGGTGTCTTTTGTTCGAAGCAGCA	6575
QY	1268	GATTACCGCGAGAAAAAAGGATCT 1292	

Db 6576 GATTACCGCGAGAAAAAAGGATCT 6600

RESULT 12
PCT-US95-08743-170/c
; Sequence 170, Application PC/TUS9508743
; GENERAL INFORMATION:
; APPLICANT:
; TITLE OF INVENTION: HUMAN NEUTRALIZING MONOCLONAL ANTIBODIES
; TITLE OF INVENTION: TO HUMAN IMMUNODEFICIENCY VIRUS
; NUMBER OF SEQUENCES: 170
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25 (EPO)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: PCT/US95/08743
; FILING DATE: 11-JUL-1995
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/276,852
; FILING DATE: 18-JUL-1994
; INFORMATION FOR SEQ ID NO: 170:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 13254 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: circular
; MOLECULE TYPE: DNA (genomic)
; PCT-US95-08743-170

Query Match 39.1%; Score 604.2; DB 5; Length 13254;
Best Local Similarity 97.9%; Pred. No. 1e-168;
Matches 612; Conservative 0; Mismatches 13; Indels 0; Gaps 0;

QY	668	TGAGAAAGATTCTCGCAGCGCGCGTGGTGGCTTTTCCATAGGCTCGCCGCC	727
Db	7279	TAGGGAGAGCGTCGACCTCGGCGCGGTGCTGGCGTTTTTCCATAGGCTCGCCGCC	7220
QY	728	TGACGAGCATCACAAAATCGAGCTCAAGTCAGAGGTGGCGAAACCCGACAGACTATA	787
Db	7219	TGACGAGCATCACAAAATCGAGCTCAAGTCAGAGGTGGCGAAACCCGACAGACTATA	7160
QY	788	AAGATACCGAGCGTTCCTCCCTCGGAAGCTCCCTCGCGCTCTCTGTCGACCTGCC	847
Db	7159	AAGATACCGAGCGTTCCTCCCTCGGAAGCTCCCTCGCGCTCTCTGTCGACCTGCC	7100
QY	848	GCTTACCGGATACCTGTCCGCTTCTCCCTCGGAAGCTGGCGCTTCTCAATGCTC	907
Db	7099	GCTTACCGGATACCTGTCCGCTTCTCCCTCGGAAGCTGGCGCTTCTCAATGCTC	7040
QY	908	ACGCTGTAGTATCTCAGTTCGGTGTAGTGTGCTTCCAAAGCTGGGCTGTGTGCACGA	967
Db	7039	ACGCTGTAGTATCTCAGTTCGGTGTAGTGTGCTTCCAAAGCTGGGCTGTGTGCACGA	6980
QY	968	ACCCCGGTTACGCGGACCGCTGGCGCTTATCCGCTAAGTATCGTCTTGAAGTCCAAACC	1027
Db	6979	ACCCCGGTTACGCGGACCGCTGGCGCTTATCCGCTAAGTATCGTCTTGAAGTCCAAACC	6920
QY	1028	GGTAAGACAGCACTATTCGCACTGGCAGCAGCCACTGGTAACAGGATTAGCAGAGCGAG	1087
Db	6919	GGTAAGACAGCACTATTCGCACTGGCAGCAGCCACTGGTAACAGGATTAGCAGAGCGAG	6860
QY	1088	GTATGTAGGCGGTGTACAGAGTTCCTTGAAGTGGTGGCTTAACCTACGGCTACACTAGAAG	1147
Db	6859	GTATGTAGGCGGTGTACAGAGTTCCTTGAAGTGGTGGCTTAACCTACGGCTACACTAGAAG	6800
QY	1148	GACAGTATTTGGTATCTGCTGAGCCAGTACCTTCCGAAAGAGATTGGTAG	1207
Db	6799	GACAGTATTTGGTATCTGCTGAGCCAGTACCTTCCGAAAGAGATTGGTAG	6740
QY	1208	CTCTTGATCCGCAACAAACACCGCTGCTAGCGGTGGTGTCTTTTGTTCGAAGCAGCA	1267


```

; SEQUENCE CHARACTERISTICS:
; LENGTH: 1905 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; US-08-906-957-9

Query Match          39.0%; Score 603; DB 2; Length 1905;
Best Local Similarity 100.0%; Pred. No. 9.5e-169;
Matches 603; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 690 GCCGCGTGTGCTGGCGGTTTTCATAGGCTCGGCCCGCTGACGAGCATCACAATAATCGA 749
DB 67 GCCGCGTGTGCTGGCGGTTTTCATAGGCTCGGCCCGCTGACGAGCATCACAATAATCGA 126

QY 750 CGCTCAAGTCAGAGGTGGGAAACCCGACAGGACTATAAAGATACACAGGCTTTCCCGCT 809
DB 127 CGCTCAAGTCAGAGGTGGGAAACCCGACAGGACTATAAAGATACACAGGCTTTCCCGCT 186

QY 810 GGAAGCTCCCTCGTGGCGCTCTCTGTTCCGACCTGCGGCTTACCGGATACCTGTCCGCC 869
DB 187 GGAAGCTCCCTCGTGGCGCTCTCTGTTCCGACCTGCGGCTTACCGGATACCTGTCCGCC 246

QY 870 TTTCTCCCTTCGGGAAGCGTGGCGGCTTTCTCAATGCTCAGCGCTAGGTATCTCAGTTGCG 929
DB 247 TTTCTCCCTTCGGGAAGCGTGGCGGCTTTCTCAATGCTCAGCGCTAGGTATCTCAGTTGCG 306

QY 930 GTGTAGTGTGCTGCTCCAAAGCTGGGCTGTGTGACGAAACCCCGCTTACGCGGACCGC 989
DB 307 GTGTAGTGTGCTGCTCCAAAGCTGGGCTGTGTGACGAAACCCCGCTTACGCGGACCGC 366

QY 990 TGGCGCTTATCCGCTAACTACGCTTTCAGTCCAAACCCGCTTAAGACAGGACTTATCGCCA 1049
DB 367 TGGCGCTTATCCGCTAACTACGCTTTCAGTCCAAACCCGCTTAAGACAGGACTTATCGCCA 426

QY 1050 CTGCGACGACGCTGTTAAGGATTTAGCAGAGGAGTATGTAGGCGGTGCTACAGAG 1109
DB 427 CTGCGACGACGCTGTTAAGGATTTAGCAGAGGAGTATGTAGGCGGTGCTACAGAG 486

QY 1110 TTTCTGAAGTGGTGGCTTAAGTACGCGTACACTAGAGGAGGATTTTGGTATCTGCGCT 1169
DB 487 TTTCTGAAGTGGTGGCTTAAGTACGCGTACACTAGAGGAGGATTTTGGTATCTGCGCT 546

QY 1170 CTGCTGAAGCCAGTTACCTTCGGAAGGAGTTCGCTGCTTTCATCCGCAACCAACC 1229
DB 547 CTGCTGAAGCCAGTTACCTTCGGAAGGAGTTCGCTGCTTTCATCCGCAACCAACC 606

QY 1230 ACCGCTGCTAGCGGTGTTTTTTTGTGGAAGCAGGAGTTACGCGCAGAAAAAGGA 1289
DB 607 ACCGCTGCTAGCGGTGTTTTTTTGTGGAAGCAGGAGTTACGCGCAGAAAAAGGA 666

QY 1290 TCT 1292
DB 667 TCT 669

RESULT 15
US-08-941-647A-1/c
; Sequence 1, Application US/08941647A
; Patent No. 5952211
; GENERAL INFORMATION:
; APPLICANT: NAKAMURA, ATSUDO
; APPLICANT: TANAKA, TOSHIO
; APPLICANT: MATSUO, YUSHI
; APPLICANT: TANASE, SUMIO
; APPLICANT: FUNATSU, MASAHICO
; APPLICANT: ETO, AKIRA
; TITLE OF INVENTION: METHOD FOR PRODUCING ACTIVE HUMAN
; NUMBER OF SEQUENCES: 4
; CORRESPONDENCE ADDRESS:

```

```

; ADDRESSEE: OBLON, SPIVAK, MCCLELLAND, MAIER & NEUSTADT,
; ADDRESSEE: P.C.
; STREET: 1755 S. JEFFERSON DAVIS HIGHWAY, SUITE 400
; CITY: ARLINGTON
; STATE: VA
; COUNTRY: USA
; ZIP: 22202
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/941,647A
; FILING DATE: 30-SEP-1997
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/533,717
; FILING DATE: 26-SEP-1995
; APPLICATION NUMBER: JP 6-268119
; FILING DATE: 07-OCT-1994
; ATTORNEY/AGENT INFORMATION:
; NAME: OBLON, NORMAN F.
; REGISTRATION NUMBER: 24,618
; REFERENCE/DOCKET NUMBER: 7335-001-0
; TELEPHONE: 703-413-3000
; TELEFAX: 703-413-2220
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2927 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; US-08-941-647A-1

Query Match          39.0%; Score 603; DB 2; Length 2927;
Best Local Similarity 100.0%; Pred. No. 1.2e-168;
Matches 603; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 690 GCCGCGTGTGCTGGCGGTTTTCATAGGCTCGGCCCGCTGACGAGCATCACAATAATCGA 749
DB 2763 GCCGCGTGTGCTGGCGGTTTTCATAGGCTCGGCCCGCTGACGAGCATCACAATAATCGA 2704

QY 750 CGCTCAAGTCAGAGGTGGGAAACCCGACAGGACTATAAAGATACACAGGCTTTCCCGCT 809
DB 2703 CGCTCAAGTCAGAGGTGGGAAACCCGACAGGACTATAAAGATACACAGGCTTTCCCGCT 2644

QY 810 GGAAGCTCCCTCGTGGCGCTCTCTGTTCCGACCTGCGGCTTACCGGATACCTGTCCGCC 869
DB 2643 GGAAGCTCCCTCGTGGCGCTCTCTGTTCCGACCTGCGGCTTACCGGATACCTGTCCGCC 2584

QY 870 TTTCTCCCTTCGGGAAGCGTGGCGGCTTTCTCAATGCTCAGCTGTAGGTATCTCAGTTGCG 929
DB 2583 TTTCTCCCTTCGGGAAGCGTGGCGGCTTTCTCAATGCTCAGCTGTAGGTATCTCAGTTGCG 2524

QY 930 GTGTAGGTCGTTCCGCTCCAAGCTGGGCTGTGTGACGAAACCCCGCTTACCGCGGACCGC 989
DB 2523 GTGTAGGTCGTTCCGCTCCAAGCTGGGCTGTGTGACGAAACCCCGCTTACCGCGGACCGC 2464

QY 990 TGGCGCTTATCCGCTAACTACGCTTTCAGTCCAAACCCGCTTAAGACAGGACTTATCGCCA 1049
DB 2463 TGGCGCTTATCCGCTAACTACGCTTTCAGTCCAAACCCGCTTAAGACAGGACTTATCGCCA 2404

QY 1050 CTGCGACGACGCTGTTAAGGATTTAGCAGAGGAGTATGTAGGCGGTGCTACAGAG 1109
DB 2403 CTGCGACGACGCTGTTAAGGATTTAGCAGAGGAGTATGTAGGCGGTGCTACAGAG 2344

QY 1110 TTTCTGAAGTGGTGGCTTAAGTACGCGTACACTAGAGGAGGATTTTGGTATCTGCGCT 1169
DB 2343 TTTCTGAAGTGGTGGCTTAAGTACGCGTACACTAGAGGAGGATTTTGGTATCTGCGCT 2284

```

Fri Jan 18 08:27:44 2002

QY 1170 CTGCTGAAGCCAGTTACCTTCGGAAGAGAGTTGCTAGCTCTTGATCCGGCAAAACAACC 1229
DB 2283 CTGCTGAAGCCAGTTACCTTCGGAAGAGAGTTGCTAGCTCTTGATCCGGCAAAACAACC 2224
QY 1230 ACGCTGTAGCGGTGTTTCTTTCAGACAGCAGATTACCGCAGAAAAAGGA 1289
DB 2223 ACGCTGTAGCGGTGTTTCTTTCAGACAGCAGATTACCGCAGAAAAAGGA 2164
QY 1290 TCT 1292
DB 2163 TCT 2161

Search completed: January 17, 2002, 11:55:46
Job time: 17473 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: January 17, 2002, 15:13:31 ; Search time 11375.6 Seconds
(without alignments)
2383.471 Million cell updates/sec

Title: US-09-242-202A-27
Perfect score: 1547
Sequence: 1 GGTACCTGCCACCATGGCGC.....CTGCACAGCCTCTCCACACA 1547

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 17159718 seqs, 8763200856 residues

Total number of hits satisfying chosen parameters: 34319436

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

Pending_Patents_NA_Main:*

1: /cgn2_6/ptodata/2/pna/pctus_COMB.seq:*

2: /cgn2_6/ptodata/2/pna/us06_COMB.seq:*

3: /cgn2_6/ptodata/2/pna/us07_COMB.seq:*

4: /cgn2_6/ptodata/2/pna/us080_COMB.seq:*

5: /cgn2_6/ptodata/2/pna/us081_COMB.seq:*

6: /cgn2_6/ptodata/2/pna/us082_COMB.seq:*

7: /cgn2_6/ptodata/2/pna/us083_COMB.seq:*

8: /cgn2_6/ptodata/2/pna/us084_COMB.seq:*

9: /cgn2_6/ptodata/2/pna/us085_COMB.seq:*

10: /cgn2_6/ptodata/2/pna/us086_COMB.seq:*

11: /cgn2_6/ptodata/2/pna/us087_COMB.seq:*

12: /cgn2_6/ptodata/2/pna/us088_COMB.seq:*

13: /cgn2_6/ptodata/2/pna/us089_COMB.seq:*

14: /cgn2_6/ptodata/2/pna/us090_COMB.seq:*

15: /cgn2_6/ptodata/2/pna/us091_COMB.seq:*

16: /cgn2_6/ptodata/2/pna/us092_COMB.seq:*

17: /cgn2_6/ptodata/2/pna/us093_COMB.seq:*

18: /cgn2_6/ptodata/2/pna/us094_COMB.seq:*

19: /cgn2_6/ptodata/2/pna/us095A_COMB.seq:*

20: /cgn2_6/ptodata/2/pna/us095B_COMB.seq:*

21: /cgn2_6/ptodata/2/pna/us095C_COMB.seq:*

22: /cgn2_6/ptodata/2/pna/us095D_COMB.seq:*

23: /cgn2_6/ptodata/2/pna/us096A_COMB.seq:*

24: /cgn2_6/ptodata/2/pna/us096B_COMB.seq:*

25: /cgn2_6/ptodata/2/pna/us096C_COMB.seq:*

26: /cgn2_6/ptodata/2/pna/us096D_COMB.seq:*

27: /cgn2_6/ptodata/2/pna/us096E_COMB.seq:*

28: /cgn2_6/ptodata/2/pna/us097A_COMB.seq:*

29: /cgn2_6/ptodata/2/pna/us097B_COMB.seq:*

30: /cgn2_6/ptodata/2/pna/us097C_COMB.seq:*

31: /cgn2_6/ptodata/2/pna/us098_COMB.seq:*

32: /cgn2_6/ptodata/2/pna/us099_COMB.seq:*

33: /cgn2_6/ptodata/2/pna/us000_COMB.seq:*

34: /cgn2_6/ptodata/2/pna/us001_COMB.seq:*

35: /cgn2_6/ptodata/2/pna/us002_COMB.seq:*

36: /cgn2_6/ptodata/2/pna/us003_COMB.seq:*

37: /cgn2_6/ptodata/2/pna/us004_COMB.seq:*

38: /cgn2_6/ptodata/2/pna/us005_COMB.seq:*

39: /cgn2_6/ptodata/2/pna/us006_COMB.seq:*

40: /cgn2_6/ptodata/2/pna/us007_COMB.seq:*

41: /cgn2_6/ptodata/2/pna/us008_COMB.seq:*

42: /cgn2_6/ptodata/2/pna/us009_COMB.seq:*

43: /cgn2_6/ptodata/2/pna/us010_COMB.seq:*

44: /cgn2_6/ptodata/2/pna/us6011_COMB.seq:*

45: /cgn2_6/ptodata/2/pna/us6012_COMB.seq:*

46: /cgn2_6/ptodata/2/pna/us6013_COMB.seq:*

47: /cgn2_6/ptodata/2/pna/us6014_COMB.seq:*

48: /cgn2_6/ptodata/2/pna/us6015_COMB.seq:*

49: /cgn2_6/ptodata/2/pna/us6016_COMB.seq:*

50: /cgn2_6/ptodata/2/pna/us6017_COMB.seq:*

51: /cgn2_6/ptodata/2/pna/us6018_COMB.seq:*

52: /cgn2_6/ptodata/2/pna/us6019_COMB.seq:*

53: /cgn2_6/ptodata/2/pna/us6020_COMB.seq:*

54: /cgn2_6/ptodata/2/pna/us6021_COMB.seq:*

55: /cgn2_6/ptodata/2/pna/us6022_COMB.seq:*

56: /cgn2_6/ptodata/2/pna/us6023_COMB.seq:*

57: /cgn2_6/ptodata/2/pna/us6024_COMB.seq:*

58: /cgn2_6/ptodata/2/pna/us6025_COMB.seq:*

59: /cgn2_6/ptodata/2/pna/us6026_COMB.seq:*

60: /cgn2_6/ptodata/2/pna/us6027_COMB.seq:*

61: /cgn2_6/ptodata/2/pna/us6028_COMB.seq:*

62: /cgn2_6/ptodata/2/pna/us6029_COMB.seq:*

63: /cgn2_6/ptodata/2/pna/us6030_COMB.seq:*

64: /cgn2_6/ptodata/2/pna/us6031_COMB.seq:*

65: /cgn2_6/ptodata/2/pna/us6032_COMB.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB	ID	Description
1	1547	100.0	1547	16	US-09-242-202-27	Sequence 27, Appl
2	1547	100.0	1547	16	US-09-242-202A-27	Sequence 27, Appl
3	1310.4	84.7	1807	16	US-09-242-202-28	Sequence 28, Appl
4	1310.4	84.7	1807	16	US-09-242-202A-28	Sequence 28, Appl
5	1086.6	70.2	2308	16	US-09-242-202-29	Sequence 29, Appl
6	1086.6	70.2	2308	16	US-09-242-202A-29	Sequence 29, Appl
7	607.4	39.3	7252	16	US-09-238-356-27	Sequence 27, Appl
8	604.2	39.1	7073	8	US-08-480-120-15	Sequence 15, Appl
9	604.2	39.1	7864	8	US-08-480-120-20	Sequence 20, Appl
10	603	39.0	608	16	US-09-242-202-26	Sequence 26, Appl
11	603	39.0	2077	1	PCT-US99-06742-7	Sequence 7, Appl
12	603	39.0	2462	18	US-09-496-445-5	Sequence 5, Appl
13	603	39.0	2462	44	US-60-118-478-5	Sequence 5, Appl
14	603	39.0	2577	3	US-07-919-535C-26	Sequence 26, Appl
15	603	39.0	2704	24	US-09-623-551-27	Sequence 27, Appl
16	603	39.0	2754	5	US-08-114-691-7	Sequence 7, Appl
17	603	39.0	2927	9	US-08-533-717-1	Sequence 1, Appl
18	603	39.0	2927	9	US-08-533-717A-1	Sequence 1, Appl
19	603	39.0	2927	15	US-09-142-481-14	Sequence 14, Appl
20	603	39.0	2939	1	PCT-US94-10146-2	Sequence 2, Appl
21	603	39.0	2939	8	US-08-488-015-2	Sequence 2, Appl
22	603	39.0	2939	15	US-09-114-835-2	Sequence 2, Appl
23	603	39.0	3175	1	PCT-US99-06704-2	Sequence 2, Appl
24	603	39.0	3175	25	US-09-646-075-2	Sequence 2, Appl
25	603	39.0	3301	8	US-08-447-430-42	Sequence 42, Appl
26	603	39.0	3301	17	US-09-342-673-42	Sequence 42, Appl
27	603	39.0	3301	21	US-09-546-411A-56	Sequence 56, Appl
28	603	39.0	3327	21	US-09-547-939A-56	Sequence 56, Appl
29	603	39.0	3327	24	US-09-628-730-56	Sequence 56, Appl
30	603	39.0	3327	45	US-60-128-830-56	Sequence 56, Appl
31	603	39.0	3331	15	US-09-148-920-1	Sequence 1, Appl
32	603	39.0	3331	18	US-09-422-565-8	Sequence 8, Appl
33	603	39.0	3344	21	US-09-546-411A-48	Sequence 48, Appl
34	603	39.0	3369	21	US-09-547-939A-48	Sequence 48, Appl
35	603	39.0	3369	24	US-09-628-730-48	Sequence 48, Appl
36	603	39.0	3369	45	US-60-128-830-48	Sequence 48, Appl
37	603	39.0	3420	25	US-09-643-192-12	Sequence 12, Appl
38	603	39.0	3421	15	US-09-130-115-31	Sequence 31, Appl
39	603	39.0	3423	8	US-08-447-430-40	Sequence 40, Appl
40	603	39.0	3423	17	US-09-342-673-40	Sequence 40, Appl
41	603	39.0	3423	17	US-09-342-673-40	Sequence 40, Appl

c 42 603 39.0 3423 18 US-09-485-737A-86 Sequence 86, Appl
c 43 603 39.0 3423 18 US-09-485-737B-86 Sequence 86, Appl
44 603 39.0 3440 3 US-07-671-058-7 Sequence 7, Appl
45 603 39.0 3462 29 US-09-742-373-10 Sequence 10, Appl

ALIGNMENTS

RESULT 1
US-09-242-202-27
; Sequence 27, Application US/09242202
; GENERAL INFORMATION:
; APPLICANT: Nelson, Edward L.
; Nelson, Peter J.
; TITLE OF INVENTION: NOVEL VECTOR FOR
; POLYNUCLEOTIDE VACCINES
; NUMBER OF SEQUENCES: 29
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: MORGAN & FINNEGAN, L.L.P.
; STREET: 345 PARK AVENUE
; CITY: NEW YORK
; STATE: NEW YORK
; COUNTRY: USA
; ZIP: 10154
; COMPUTER READABLE FORM:
; MEDIUM TYPE: DISKETTE, 3.5 INCH, 1.44 MB
; COMPUTER: IBM PC COMPATIBLE
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: MS WORD 97
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/242,202
; FILING DATE: 01-Nov-1999
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US60/023931
; FILING DATE: 14-AUG-1996
; ATTORNEY/AGENT INFORMATION:
; NAME: KATHRYN M. BROWN
; REGISTRATION NUMBER: 34556
; REFERENCE/DOCKET NUMBER: 2026-4236US1
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 758-4800
; TELEFAX: (212) 751-6849
; TELEX: 421792
; INFORMATION FOR SEQ ID NO: 27:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1547 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: unknown
; MOLECULE TYPE: cDNA
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
; SEQUENCE DESCRIPTION: SEQ ID NO: 27:

Query Match 100.0%; Score 1547; DB 16; Length 1547;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1547; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGTACCTGCCACCATGGCGCGGATTTCTTATCACTGATAAGTTGGTGACATATTATGTT 60
DB 1 GGTACCTGCCACCATGGCGCGGATTTCTTATCACTGATAAGTTGGTGACATATTATGTT 60
QY 61 TATCAGTGATAAGTCTCAAGCATGACAAAGTTGACGCCGATAACATGATCGTCCCGG 120
DB 61 TATCAGTGATAAGTCTCAAGCATGACAAAGTTGACGCCGATAACATGATCGTCCCGG 120
QY 121 CCCTGGACTGTTGAACGAGGTGCGCGTAGACGCTCTGACGACACGCAAACTGCGCGAAG 180
DB 121 CCCTGGACTGTTGAACGAGGTGCGCGTAGACGCTCTGACGACACGCAAACTGCGCGAAG 180

QY 181 GTTGGGGTGCAGCAGCGCGCTTTTACTGGCACTTCAGGAACAAGCGGCGCCCTTAAGG 240
DB 181 GTTGGGGTGCAGCAGCGCGCTTTTACTGGCACTTCAGGAACAAGCGGCGCCCTTAAGG 240
QY 241 GCCATATGTTGAGTGGATGCTTGACCCAGCGGGGATGGGGGAGACCTGTAGTCAGAG 300
DB 241 GCCATATGTTGAGTGGATGCTTGACCCAGCGGGGATGGGGGAGACCTGTAGTCAGAG 300
QY 301 CCCCGGCGCAGCACAGGCCAATGCCGCTCTTCCCTGCAGGATGAGTAGTGCCCTC 360
DB 301 CCCCGGCGCAGCACAGGCCAATGCCGCTCTTCCCTGCAGGATGAGTAGTGCCCTC 360
QY 361 TCCTGGCCCTTGAAGTTGCCACTCCAGTGCACACGACCTTCTCCTAATAAATAAATTAAGTT 420
DB 361 TCCTGGCCCTTGAAGTTGCCACTCCAGTGCACACGACCTTCTCCTAATAAATAAATTAAGTT 420
QY 421 GCATCATTTTGTCTGACTAGTGTCTCTATATATATAAGCTTGATATCGAATCTTT 480
DB 421 GCATCATTTTGTCTGACTAGTGTCTCTATATATATAAGCTTGATATCGAATCTTT 480
QY 481 CGGACTTTTGAAGTGATGTTGGTGGGGAAGGATTCGAACCTTCGAAGTCGATGACGCG 540
DB 481 CGGACTTTTGAAGTGATGTTGGTGGGGAAGGATTCGAACCTTCGAAGTCGATGACGCG 540
QY 541 AGATTTAGAGTCTGCTCTCTTTGGCGGCTCGGGAACCCACACGCGGTAATGCTTTTACT 600
DB 541 AGATTTAGAGTCTGCTCTCTTTGGCGGCTCGGGAACCCACACGCGGTAATGCTTTTACT 600
QY 601 GCCTGCTCCCTTATCGGAAGCGGGCGCATCATATCAATGACGCGCGCTGTAAGT 660
DB 601 GCCTGCTCCCTTATCGGAAGCGGGCGCATCATATCAATGACGCGCGCTGTAAGT 660
QY 661 GTTAGCTTGAGAAAGATTCCTGACGCGCGCGCTTGCTGGCGTTTTCATAGGCTCC 720
DB 661 GTTAGCTTGAGAAAGATTCCTGACGCGCGCGCTTGCTGGCGTTTTCATAGGCTCC 720
QY 721 GCCCGCTGCAGGATCACAAAATCGAGCTCAAGTCAGAGTGGCGCAACCCACAG 780
DB 721 GCCCGCTGCAGGATCACAAAATCGAGCTCAAGTCAGAGTGGCGCAACCCACAG 780
QY 781 GACTATAAGATACCAAGCGGTTTCCCGCTGGAAGCTCCCTGCGCTCTCTCTGTTCCGA 840
DB 781 GACTATAAGATACCAAGCGGTTTCCCGCTGGAAGCTCCCTGCGCTCTCTCTGTTCCGA 840
QY 841 CCCTGCGCTTACCGGATACCTGTCGCGCTTCTCCTTCGGGAAGCGTGGCGCTTCTC 900
DB 841 CCCTGCGCTTACCGGATACCTGTCGCGCTTCTCCTTCGGGAAGCGTGGCGCTTCTC 900
QY 901 AATGCTCAGCTGTAGGTATCTCAGTTCGCTAGTGTGCTCGCTCCAAAGCTGGGCTGTG 960
DB 901 AATGCTCAGCTGTAGGTATCTCAGTTCGCTAGTGTGCTCGCTCCAAAGCTGGGCTGTG 960
QY 961 TGCACGAACCCCGCTTACGCCGCGCTGCGCTTATCCGGTAACTATCGTCTTTGAGT 1020
DB 961 TGCACGAACCCCGCTTACGCCGCGCTGCGCTTATCCGGTAACTATCGTCTTTGAGT 1020
QY 1021 CCAACCCGCTAAGACACGACTTATCCCGACTGCGACGACGCTGCTGTAACAGGATTAGCA 1080
DB 1021 CCAACCCGCTAAGACACGACTTATCCCGACTGCGACGACGCTGCTGTAACAGGATTAGCA 1080
QY 1081 GAGCGAGGTATGTAGCGGCTGCTACAGAGTCTTGAAGTGGTGGCCCTAACTACGGGTACA 1140
DB 1081 GAGCGAGGTATGTAGCGGCTGCTACAGAGTCTTGAAGTGGTGGCCCTAACTACGGGTACA 1140
QY 1141 CTAGAGGACAGTATTTGGTATCTGCGCTCTGCTGAAGCCAGTTACTCTCGGAAAAAGAG 1200
DB 1141 CTAGAGGACAGTATTTGGTATCTGCGCTCTGCTGAAGCCAGTTACTCTCGGAAAAAGAG 1200
QY 1201 TTGGTAGCTCTTGATCCGGCAACAAACACCGCTGGTAGCGGTGTTTTTTTGTGTGCA 1260
DB 1201 TTGGTAGCTCTTGATCCGGCAACAAACACCGCTGGTAGCGGTGTTTTTTTGTGTGCA 1260
QY 1261 AGCAGCAGATTACGCCGCAAGAAAAAGGATCTGCGGGATCCCGGAGAGCTCACTCTAGATG 1320

Db 1261 AGCAGCAGATTAGCCGCAAAAAAGGATCTGGGGATCCGGAGAGCTCACTCTAGATG 1320
QY 1321 AGAGCAGCTGAGGGAGACAGAGACTCGAATTTCCGGAGCTATTTCAGTTTCTTTTC 1380
Db 1321 AGAGCAGCTGAGGGAGACAGAGACTCGAATTTCCGGAGCTATTTCAGTTTCTTTTC 1380
QY 1381 CGTTTGTCAATTTCACTTATGATACCGGCCAATCTTGGTTGCTATTTTGGAAACTCC 1440
Db 1381 CGTTTGTCAATTTCACTTATGATACCGGCCAATCTTGGTTGCTATTTTGGAAACTCC 1440
QY 1441 CCTTAGGGATGCCCTCAACTGGCCCTATAAAGGCCAGCCTGAGCTGCAGAGGATTC 1500
Db 1441 CCTTAGGGATGCCCTCAACTGGCCCTATAAAGGCCAGCCTGAGCTGCAGAGGATTC 1500
QY 1501 TGCAGAGGATCAAGACAGCAGCTGGACCTCGCACACCTCTCCACA 1547
Db 1501 TGCAGAGGATCAAGACAGCAGCTGGACCTCGCACACCTCTCCACA 1547

RESULT 2
US-09-242-202A-27
; Sequence 27, Application US/09242202A
; GENERAL INFORMATION:
; APPLICANT: Nelson, Edward L.
; Nelson, Peter J.
; TITLE OF INVENTION: NOVEL VECTOR FOR
; POLYNUCLEOTIDE VACCINES
; NUMBER OF SEQUENCES: 29
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: MORGAN & FINNEGAN, L.L.P.
; STREET: 345 PARK AVENUE
; CITY: NEW YORK
; STATE: NEW YORK
; COUNTRY: USA
; ZIP: 10154
; COMPUTER READABLE FORM:
; MEDIUM TYPE: DISKETTE, 3.5 INCH, 1.44 MB
; COMPUTER: IBM PC COMPATIBLE
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: MS WORD 97
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/242, 202A
; FILING DATE: 20-Apr-2000
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US60/023931
; FILING DATE: 14-AUG-1996
; ATTORNEY/AGENT INFORMATION:
; NAME: KATHRYN M. BROWN
; REGISTRATION NUMBER: 34556
; REFERENCE/DOCKET NUMBER: 2026-4236US1
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 758-4800
; TELEFAX: (212) 751-6849
; TELEX: 421792
; INFORMATION FOR SEQ ID NO: 27:
; SEQUENCE CHARACTERISTICS:
; TYPE: nucleic acid
; LENGTH: 1547 base pairs
; STRANDEDNESS: single
; TOPOLOGY: unknown
; MOLECULE TYPE: cDNA
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
; SEQUENCE DESCRIPTION: SEQ ID NO: 27:
US-09-242-202A-27

Query Match 100.0%; Score 1547; DB 16; Length 1547;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1547; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 GGTACCTGCACCATGGCGGGATTCTTTATCACTGATAAGTTGGTGGACATATTATGTT 60

Db 1 GGTACCTGCACCATGGCGGGATTCTTTATCACTGATAAGTTGGTGGACATATTATGTT 60
QY 61 TATCAGTGATAAAGTGTCAAGCATGACAAAGTTGCAGCCGAATACAGTGCCTGCCGG 120
Db 61 TATCAGTGATAAAGTGTCAAGCATGACAAAGTTGCAGCCGAATACAGTGCCTGCCGG 120
QY 121 CCCTGGACTGTTGAACGAGGTGCGCGCTAGACGCTGTGACGACACGCAAACTGGCGAAACG 180
Db 121 CCCTGGACTGTTGAACGAGGTGCGCGCTAGACGCTGTGACGACACGCAAACTGGCGAAACG 180
QY 181 GTTGGGGGTGCACAGCGCGCTTTACTGGCACTTTCAGGAAACAAGCGGCGCTTTAAGG 240
Db 181 GTTGGGGGTGCACAGCGCGCTTTACTGGCACTTTCAGGAAACAAGCGGCGCTTTAAGG 240
QY 241 GCCATATGGTGTAGTGGATGCTTACCCAGCGGGGATGGGGAGACCTGTAGTCAGAG 300
Db 241 GCCATATGGTGTAGTGGATGCTTACCCAGCGGGGATGGGGAGACCTGTAGTCAGAG 300
QY 301 CCCCCGGGACACAGGCCAATGCCGCTCCTTCCCTTCAGAGGATGAGTAGTGCTC 360
Db 301 CCCCCGGGACACAGGCCAATGCCGCTCCTTCCCTTCAGAGGATGAGTAGTGCTC 360
QY 361 TCCTGGCCCTGGAAAGTTGCCACTCCAGTGCCCAAGCCTTGTCTTAATAAATAAAGTT 420
Db 361 TCCTGGCCCTGGAAAGTTGCCACTCCAGTGCCCAAGCCTTGTCTTAATAAATAAAGTT 420
QY 421 GCATCATTTTGTCTGACTAGGTGTCTCTATAATAATAAAGCTTGATATCGAATCTTT 480
Db 421 GCATCATTTTGTCTGACTAGGTGTCTCTATAATAATAAAGCTTGATATCGAATCTTT 480
QY 481 CGGACTTTTGAAGAGTGTGGTGGGGAGGATTCGAACCTTCGAAGTCGATGACGGC 540
Db 481 CGGACTTTTGAAGAGTGTGGTGGGGAGGATTCGAACCTTCGAAGTCGATGACGGC 540
QY 541 AGATTTAGAGTCTGCTCCCTTTGGCGCTCGGAAACCCACCGGTAATGCTTTTACT 600
Db 541 AGATTTAGAGTCTGCTCCCTTTGGCGCTCGGAAACCCACCGGTAATGCTTTTACT 600
QY 601 GGCTGCTCCTTATCGGGAAGCGGGCGCATCATATCAAAATGACGCGCTGTAAGT 660
Db 601 GGCTGCTCCTTATCGGGAAGCGGGCGCATCATATCAAAATGACGCGCTGTAAGT 660
QY 661 GTTACGTTGAGAAAGATTCCTGCAGCCCGCGCTTGTGGCGTTTTCATAGGCTCC 720
Db 661 GTTACGTTGAGAAAGATTCCTGCAGCCCGCGCTTGTGGCGTTTTCATAGGCTCC 720
QY 721 GCCCCCTGACGAGCATCAGCAAAATCGACGCTCAAGTCAGAGGTGGCGAAACCCGACAG 780
Db 721 GCCCCCTGACGAGCATCAGCAAAATCGACGCTCAAGTCAGAGGTGGCGAAACCCGACAG 780
QY 781 GACTATAAAGATACAGGCGGTTCCCGCTGGAAGCTCCCTGCGCTCTCTGTTCCGA 840
Db 781 GACTATAAAGATACAGGCGGTTCCCGCTGGAAGCTCCCTGCGCTCTCTGTTCCGA 840
QY 841 CCCTGCGCTTACCGGATACCTGTCCGCTTCTCCCTCGGGAAGCGTGGCGTTTCTC 900
Db 841 CCCTGCGCTTACCGGATACCTGTCCGCTTCTCCCTCGGGAAGCGTGGCGTTTCTC 900
QY 901 AATGCTCACGCTGTAGGTATCTCAGTTTCAGTTTCGCTGAGGTTCGCTCAAGCTGGGCTGTG 960
Db 901 AATGCTCACGCTGTAGGTATCTCAGTTTCGCTGAGGTTCGCTCAAGCTGGGCTGTG 960
QY 961 TGCACGAACCCCGCTTACGCGCGAGCGCTTATCCGTAACATATCGTCTTGAAGT 1020
Db 961 TGCACGAACCCCGCTTACGCGCGAGCGCTTATCCGTAACATATCGTCTTGAAGT 1020
QY 1021 CCAACCGGTAAAGACAGACTTATCGCACTGCGCAGCAGCACTGGTAACAGGATTAGCA 1080
Db 1021 CCAACCGGTAAAGACAGACTTATCGCACTGCGCAGCAGCACTGGTAACAGGATTAGCA 1080
QY 1081 GAGCGAGGTATGAGCGGTGCTACAGAGTTCTTGAAGTGGTGGCTTAACAGCGGTACA 1140

QY 961 TGCACGAACCCCGGTTTCAGCCGACCGCTGCGCCCTTATCCCGTAATATCGTCTTGGT 1020
Db 961 TGCACGAACCCCGGTTTCAGCCGACCGCTGCGCCCTTATCCCGTAATATCGTCTTGGT 1020
QY 1021 CCACCCGGTAGACACGACTTATCCCACTGCGCAGCAGCAGCTGTATACAGATTAGCA 1080
Db 1021 CCACCCGGTAGACACGACTTATCCCACTGCGCAGCAGCAGCTGTATACAGATTAGCA 1080
QY 1081 GAGCAGGATGATAGCGGTGCTACAGAGTTCTTGAAGTGGTGGCCCTAACTACGGGTACA 1140
Db 1081 GAGCAGGATGATAGCGGTGCTACAGAGTTCTTGAAGTGGTGGCCCTAACTACGGGTACA 1140
QY 1141 CTAGAAGGACAGATTGTTGTTATCTGCGCTCTGCTGAAGCCAGTTACTTTCGGAAGAG 1200
Db 1141 CTAGAAGGACAGATTGTTGTTATCTGCGCTCTGCTGAAGCCAGTTACTTTCGGAAGAG 1200
QY 1201 TTGGTAGCTTTCATCCGGCAACAAACCCGCTGGTAGCGGTGTTTTTTTGTGCA 1260
Db 1201 TTGGTAGCTTTCATCCGGCAACAAACCCGCTGGTAGCGGTGTTTTTTTGTGCA 1260
QY 1261 ACCAGCAGATTAGCCGACAGAAAAAGGATCTGGGGATCCGGAGAGCTCAC 1312
Db 1261 ACCAGCAGATTAGCCGACAGAAAAAGGATCTGGGGATCCGGAGAGCTCCC 1312

RESULT 4
US-09-242-202A-28
; Sequence 28, Application US/09242202A
; GENERAL INFORMATION:
; APPLICANT: Nelson, Edward L.
; APPLICANT: Nelson, Peter J.
; TITLE OF INVENTION: POLYNUCLEOTIDE VACCINES
; NUMBER OF SEQUENCES: 29
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: MORGAN & FINNEGAN, L.L.P.
; STREET: 345 PARK AVENUE
; CITY: NEW YORK
; STATE: NEW YORK
; COUNTRY: USA
; ZIP: 10154
; COMPUTER READABLE FORM:
; MEDIUM TYPE: DISKETTE, 3.5 INCH, 1.44 MB
; COMPUTER: IBM PC COMPATIBLE
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: MS WORD 97
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/242,202A
; FILING DATE: 20-Apr-2000
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US60/023931
; FILING DATE: 14-AUG-1996
; ATTORNEY/AGENT INFORMATION:
; NAME: KATHRYN M. BROWN
; REGISTRATION NUMBER: 34556
; REFERENCE/DOCKET NUMBER: 2026-4236US1
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 758-4800
; TELEFAX: (212) 751-6849
; TELE: 421792
; INFORMATION FOR SEQ ID NO: 28:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1807 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: unknown
; MOLECULE TYPE: cDNA
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
; SEQUENCE DESCRIPTION: SEQ ID NO: 28:
US-09-242-202A-28

Query Match 84.7%; Score 1310.4; DB 16; Length 1807;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 1311; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 1 GGTACCTGCCACCATCGCGGATTTCTTATCACTGATAGTTGGTGGACATATTATGTT 60
Db 1 GGTACCTGCCACCATCGCGGATTTCTTATCACTGATAGTTGGTGGACATATTATGTT 60
QY 61 TATCAGTGTAAAGTGTCAAGCATGACAAAGTTGCAGCCGCAATACAGTATCGTGCCTG 120
Db 61 TATCAGTGTAAAGTGTCAAGCATGACAAAGTTGCAGCCGCAATACAGTATCGTGCCTG 120
QY 121 CCTTGGACTGTTGAACGAGTTCGCGGTAGACGCTCTGACACACGCAAACTGCGGGAACG 180
Db 121 CCTTGGACTGTTGAACGAGTTCGCGGTAGACGCTCTGACACACGCAAACTGCGGGAACG 180
QY 181 GTTGGGGGTGCAGCAGCCGCGCTTTACTGGCACTTCAGGAACAAGCGGCGCTTTAAGG 240
Db 181 GTTGGGGGTGCAGCAGCCGCGCTTTACTGGCACTTCAGGAACAAGCGGCGCTTTAAGG 240
QY 241 GCCATATGTTGAGTGTGATGCTTGACCCAGCGGGGATGGGGGAGACCTGTAGTCAGAG 300
Db 241 GCCATATGTTGAGTGTGATGCTTGACCCAGCGGGGATGGGGGAGACCTGTAGTCAGAG 300
QY 301 CCCCCTGGGAGCAGCAGCCCAATGCCGCTCTTCCCTGCAGGATGAGTGTAGTGCCTC 360
Db 301 CCCCCTGGGAGCAGCAGCCCAATGCCGCTCTTCCCTGCAGGATGAGTGTAGTGCCTC 360
QY 361 TCCTGGCCCTTGGAAAGTTGCCACTCCAGTGCACCCAGCCCTTGTCTTAATAAATTAAGTT 420
Db 361 TCCTGGCCCTTGGAAAGTTGCCACTCCAGTGCACCCAGCCCTTGTCTTAATAAATTAAGTT 420
QY 421 GCATCATTTTGTCTGACTAGTGTCTCTATATATATATATATATATATATATATATAT 480
Db 421 GCATCATTTTGTCTGACTAGTGTCTCTATATATATATATATATATATATATATATAT 480
QY 481 CGGACTTTTGAAGTGTGTTGGGGAAGGATTCGAACCTTCGAAGTCGATGACGGC 540
Db 481 CGGACTTTTGAAGTGTGTTGGGGAAGGATTCGAACCTTCGAAGTCGATGACGGC 540
QY 541 AGATTTAGAGTCTGCTCCCTTTGGCCGCTCGGGAACCCACCACCGGTAATGCTTTTACT 600
Db 541 AGATTTAGAGTCTGCTCCCTTTGGCCGCTCGGGAACCCACCACCGGTAATGCTTTTACT 600
QY 601 GGCTGCTCCCTTATCGGGAAGCGGGCGCATCATATCAATACAGCCGCTGTAAAGT 660
Db 601 GGCTGCTCCCTTATCGGGAAGCGGGCGCATCATATCAATACAGCCGCTGTAAAGT 660
QY 661 GTTACGTTGAGAAAGATTCCTGCAGCCCGCGCTTGTGCGGTTTTCATAGGCTCC 720
Db 661 GTTACGTTGAGAAAGATTCCTGCAGCCCGCGCTTGTGCGGTTTTCATAGGCTCC 720
QY 721 GCGCCCTGCAGCAGCATCAAAAAATCGACGCTCAAGTCAGAGTGGCGGAACCCGACAG 780
Db 721 GCGCCCTGCAGCAGCATCAAAAAATCGACGCTCAAGTCAGAGTGGCGGAACCCGACAG 780
QY 781 GACTATAAAGATACAGGCGTTTCCCTCGGAAGCTCCCTCGTCCGCTCTCTGTTCCGA 840
Db 781 GACTATAAAGATACAGGCGTTTCCCTCGGAAGCTCCCTCGTCCGCTCTCTGTTCCGA 840
QY 841 CCTTGGCGCTTACGGGATACCTGTGCGCCCTTCTCCCTCGGAAGCGTGGCGCTTCTC 900
Db 841 CCTTGGCGCTTACGGGATACCTGTGCGCCCTTCTCCCTCGGAAGCGTGGCGCTTCTC 900
QY 901 AATGCTCACGCTGTAGTATCTCAGTTGGTGTAGTGTCTCGCTCCCAAGCTGGCTGTG 960
Db 901 AATGCTCACGCTGTAGTATCTCAGTTGGTGTAGTGTCTCGCTCCCAAGCTGGCTGTG 960
QY 961 TGCACGAACCCCGGTTTCAGCCGACCGCTGCGCCCTTATCCCGTAATATCGTCTTGGT 1020
Db 961 TGCACGAACCCCGGTTTCAGCCGACCGCTGCGCCCTTATCCCGTAATATCGTCTTGGT 1020

[illegible]

```

RESULT 7
US-09-238-356-27
; Sequence 27, Application US/09238356
; GENERAL INFORMATION:
; APPLICANT: Kingsman, et al
; TITLE OF INVENTION: Retroviral Vectors
; FILE REFERENCE: 674523-2006
; CURRENT APPLICATION NUMBER: US/09/238,356
; CURRENT FILING DATE: 1999-01-27
; PRIOR APPLICATION NUMBER: PCT/CB/03876
; PRIOR FILING DATE: 1998-12-27
; NUMBER OF SEQ IDS NOS: 64
; SOFTWARE: PatentIn version 3.0

```


Db	3156	ACCCCGGTTACGCCGACCGCTGGCCCTTATCCGGTAACATATCGTCTTGAGTCCAAACC	3215
Qy	1028	GGTAAGACACGACTTATCGCCCACTGGCAGACGCCACTGGTAACAGGATTAGCAGAGCGAG	1087
Db	3216	GGTAAGACACGACTTATCGCCCACTGGCAGACGCCACTGGTAACAGGATTAGCAGAGCGAG	3275
Qy	1088	GTATGTAGCGGTGTACAGAGTCTTGAACGTGTGGCTTAACATCGGCTACACTAGAA	1147
Db	3276	GTATGTAGCGGTGTACAGAGTCTTGAACGTGTGGCTTAACATCGGCTACACTAGAA	3335
Qy	1148	GACAGTATTGGTATCTGCGCTCTCTGAAGCCAGTTTACCTTCGSAAGAGAGTTGGTAG	1207
Db	3336	GACAGTATTGGTATCTGCGCTCTCTGAAGCCAGTTTACCTTCGSAAGAGAGTTGGTAG	3395
Qy	1208	CTCTTGATCGGCAACAACCAACCGCTGGTAGCGGTGGTTTTTTTGTTCGACAGCAGCA	1267
Db	3396	CTCTTGATCGGCAACAACCAACCGCTGGTAGCGGTGGTTTTTTTGTTCGACAGCAGCA	3455
Qy	1268	GATTACGCGCAGAAAAAGGATCT	1292
Db	3456	GATTACGCGCAGAAAAAGGATCT	3480

RESULT 9
US-88-480-120-20
; Sequence 20, Application US/08480120
; GENERAL INFORMATION:
; APPLICANT: Joliffe, Linda K.
; APPLICANT: Zivin, Robert A.
; APPLICANT: Pulito, Virginia L.
; TITLE OF INVENTION: CDR-GRAFTED ANTI-TISSUE FACTOR
; TITLE OF INVENTION: ANTIBODIES AND METHODS OF USE THEREOF
; NUMBER OF SEQUENCES: 20

```

1 COMPUTER READABLE FORM:
2 MEDIUM TYPE: Floppy disk
3 COMPUTER: IBM PC compatible
4 OPERATING SYSTEM: PC-DOS/MS-DOS
5 SOFTWARE: PatentIn Release #1.0, Version #1.25
6 CURRENT APPLICATION DATA:
7 APPLICATION NUMBER: US/08/480,120
8 FILING DATE: 07-JUN-1995
9

```

Query Match	39.1%;	Score 604.2;	DB 8;	Length 7864;
Best Local Similarity	97.9%;	Pred. NO. 7.7e-167;		
Matches 612; Conservative	0;	Mismatches 13;	Indels	0;

Qy	668	TGAGAAGAAATTCCTCGAGCCCGCCGCTTGCTGGCGTTTTCATAGGCTCCGCCCC	727
Db	1255	TAAGGAGAGCGCTCGACCTCGGCCCGCGTGTGCGGTTTTTTCATAGGCTCCGCCCC	1314
Qy	728	TGACGAGCATCACAAAATTCGACGCTCAAGTCAGAGGTGGGAAACCCGACGAGCTATA	787
Db	1315	TGACGAGCATCACAAAATTCGACGCTCAAGTCAGAGGTGGGAAACCCGACGAGCTATA	1374
Qy	788	AAGATACGAGCGTTTCCCCCTGGAAAGTCCCTCGTCGGCTCTCTCTGTTCGACCCCTGCC	847
Db	1375	AAGATACGAGCGTTTCCCCCTGGAAAGTCCCTCGTCGGCTCTCTCTGTTCGACCCCTGCC	1434
Qy	848	GCATTACCGGATACCTCTCGCGCTTTCTCCCTTTCGGGAACGCTGGCGCTTCTCAATGCTC	907
Db	1435	GCATTACCGGATACCTCTCGCGCTTTCTCCCTTTCGGGAACGCTGGCGCTTCTCAATGCTC	1494
Qy	908	ACGCTGTAGGTATCTCAGTTTCGGTGTAGTCTGTGCTCTCAAGCTGGGCTGTGTGCACGA	967
Db	1495	ACGCTGTAGGTATCTCAGTTTCGGTGTAGTCTGTGCTCTCAAGCTGGGCTGTGTGCACGA	1554
Qy	968	ACCCCGCGTTACGCGGACGGCTGGCGCTTATCCGGTAACATATCGTCTGTAGTCTCAACCC	1027
Db	1555	ACCCCGCGTTACGCGGACGGCTGGCGCTTATCCGGTAACATATCGTCTGTAGTCTCAACCC	1614
Qy	1028	GGTAAGACAGCTTATCGCCACTGCGACGACGCACTGGTAACAGGATTACGACAGCGAG	1087
Db	1615	GGTAAGACAGCTTATCGCCACTGCGACGACGCACTGGTAACAGGATTACGACAGCGAG	1674
Qy	1088	GTATGTAGCGGCTGTACAGATTCCTTGAAGTGGTGGCTTAACACTACGGCTACACTAGAAG	1147
Db	1675	GTATGTAGCGGCTGTACAGATTCCTTGAAGTGGTGGCTTAACACTACGGCTACACTAGAAG	1734
Qy	1148	GACAGTATTTGGTATCTGGCGCTCTGCTGAAGCGAGTTTACCTTCGGAAGAGGTTGGTAG	1207
Db	1735	GACAGTATTTGGTATCTGGCGCTCTGCTGAAGCGAGTTTACCTTCGGAAGAGGTTGGTAG	1794
Qy	1208	CTCTTGATCCGGCAAAACAAACACCGCTGGTAGCGGTGGTTTTTTTGTTCGAAGCAGCA	1267
Db	1795	CTCTTGATCCGGCAAAACAAACACCGCTGGTAGCGGTGGTTTTTTTGTTCGAAGCAGCA	1854
Qy	1268	GATTACCGCAGAAAAAAGGATCT	1292
Db	1855	GATTACCGCAGAAAAAAGGATCT	1879

RESULT 10
US-09-242-26
: Sequence 26, Application US/09242202
: GENERAL INFORMATION:
: APPLICANT: Nelson, Edward L.
: APPLICANT: Nelson, Peter J.
: TITLE OF INVENTION: NOVEL VECTOR FOR
: POLYNUCLEOTIDE VACCINES
: NUMBER OF SEQUENCES: 29
: CORRESPONDENCE ADDRESS:
: ADDRESSEE: MORGAN & FINNEGAN, L.L.P.
: STREET: 345 PARK AVENUE
: CITY: NEW YORK
: STATE: NEW YORK
: COUNTRY: USA
: ZIP: 10154
: COMPUTER READABLE FORM:
: MEDIUM TYPE: DISKETTE, 3.5 INCH, 1.44 MB
: COMPUTER: IBM PC COMPATIBLE
: OPERATING SYSTEM: PC-DOS/MS-DOS
: SOFTWARE: MS WORD 97
: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: US/09/242,202
: FILING DATE: 01-NOV-1999
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER: US60/023931
: FILING DATE: 14-AUG-1996

ATTORNEY/AGENT INFORMATION:
NAME: KATHRYN M. BROWN
REGISTRATION NUMBER: 34556
REFERENCE/DOCKET NUMBER: 2026-42360US1
TELECOMMUNICATION INFORMATION:
TELEPHONE: (212) 758-4800
TELEFAX: (212) 751-6849
TELEX: 421792
INFORMATION FOR SEQ ID NO: 26:
SEQUENCE CHARACTERISTICS:
LENGTH: 608 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: unknown
MOLECULE TYPE: cdna
HYPOTHETICAL: NO
ANTI-SENSE: No
SEQUENCE DESCRIPTION: SEQ ID NO: 26:
US-09-242-202-26

Query Match 39.0%; Score 603; DB 16; Length 608;
Best Local Similarity 100.0%; Pred. No. 6e-167;
Matches 603; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 690 GCCGCTTGGCTGGCGTTTTCATAGGCTCCGCCCTGACGAGCATCACAAAAATCGA 749
Db |
QY 750 CGCTCAAGTCAGAGTGGCGAAACCCGACAGACTATAAAGATACAGGCGTTTCCCGCT 809
Db |
QY 66 CGCTCAAGTCAGAGTGGCGAAACCCGACAGACTATAAAGATACAGGCGTTTCCCGCT 125
Db |
QY 810 GGAAGCTCCCTGCTGGCGTCTCTCTCCGACCTGCCGCTTACCGGATACCTGTCGCC 869
Db |
QY 126 GGAAGCTCCCTGCTGGCGTCTCTCTCCGACCTGCCGCTTACCGGATACCTGTCGCC 185
Db |
QY 870 TTTCTCCCTTCCGGAAGCGTGGCGCTTCTCAATGCTACGCTAGGTATCTCAGTTGC 929
Db |
QY 186 TTTCTCCCTTCCGGAAGCGTGGCGCTTCTCAATGCTACGCTAGGTATCTCAGTTGC 245
Db |
QY 930 GTGTAGGTCGTTCCGCTCCCAAGCTGGGCTGTGTGACGAAACCCCGCTTACCGGACCG 989
Db |
QY 246 GTGTAGGTCGTTCCGCTCCCAAGCTGGGCTGTGTGACGAAACCCCGCTTACCGGACCG 305
Db |
QY 990 TGCCTTATCCGGTAACATATCGTCTTGAAGTCCAAACCCGGTAAGACACGACTTATCGCA 1049
Db |
QY 306 TGCCTTATCCGGTAACATATCGTCTTGAAGTCCAAACCCGGTAAGACACGACTTATCGCA 365
Db |
QY 1050 CTGCAGCAGCCACTGGTAACAGATTAGCAGAGCGAGGTATGTAGCGGTGCTACAGAG 1109
Db |
QY 366 CTGCAGCAGCCACTGGTAACAGATTAGCAGAGCGAGGTATGTAGCGGTGCTACAGAG 425
Db |
QY 1110 TTTCTGAAGTGGTGGCTTAACACTACGCTACACTAGAGGACAGATTATGGTATCTCGCT 1169
Db |
QY 426 TTTCTGAAGTGGTGGCTTAACACTACGCTACACTAGAGGACAGATTATGGTATCTCGCT 485
Db |
QY 1170 CTGCTGAAGCAGTACCTTCGGAAGAGTTGGTAGCTTGTATCCGCAACAAACC 1229
Db |
QY 486 CTGCTGAAGCAGTACCTTCGGAAGAGTTGGTAGCTTGTATCCGCAACAAACC 545
Db |
QY 1230 ACCGCTGGTAGCGGTGTTTTTTTGTGCAAGCAGCAGATTACGCGCAAAAAAAGGA 1289
Db |
QY 546 ACCGCTGGTAGCGGTGTTTTTTTGTGCAAGCAGCAGATTACGCGCAAAAAAAGGA 1289
Db |
QY 1290 TCT 1292
Db |
QY 606 TCT 608
Db |

RESULT 11
US-09-242-202a-26
; Sequence 26, Application US/09242202a

GENERAL INFORMATION:
APPLICANT: Nelson, Edward L.
TITLE OF INVENTION: NOVEL VECTOR FOR
POLYNUCLEOTIDE VACCINES
NUMBER OF SEQUENCES: 29
CORRESPONDENCE ADDRESS:
ADDRESSEE: MORGAN & FINNEGAN, L.L.P.
STREET: 345 PARK AVENUE
CITY: NEW YORK
STATE: NEW YORK
COUNTRY: USA
ZIP: 10154
COMPUTER READABLE FORM:
MEDIUM TYPE: DISKETTE, 3.5 INCH, 1.44 MB
COMPUTER: IBM PC COMPATIBLE
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: MS WORD 97
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/242, 202A
FILING DATE: 20-Apr-2000
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US60/023931
FILING DATE: 14-AUG-1996
ATTORNEY/AGENT INFORMATION:
NAME: KATHRYN M. BROWN
REGISTRATION NUMBER: 34556
REFERENCE/DOCKET NUMBER: 2026-42360US1
TELECOMMUNICATION INFORMATION:
TELEPHONE: (212) 758-4800
TELEFAX: (212) 751-6849
TELEX: 421792
INFORMATION FOR SEQ ID NO: 26:
SEQUENCE CHARACTERISTICS:
LENGTH: 608 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: unknown
MOLECULE TYPE: cdna
HYPOTHETICAL: NO
ANTI-SENSE: No
SEQUENCE DESCRIPTION: SEQ ID NO: 26:
US-09-242-202a-26

Query Match 39.0%; Score 603; DB 16; Length 608;
Best Local Similarity 100.0%; Pred. No. 6e-167;
Matches 603; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 690 GCCGCTTGGCTGGCGTTTTCATAGGCTCCGCCCTGACGAGCATCACAAAAATCGA 749
Db |
QY 750 CGCTCAAGTCAGAGTGGCGAAACCCGACAGACTATAAAGATACAGGCGTTTCCCGCT 809
Db |
QY 66 CGCTCAAGTCAGAGTGGCGAAACCCGACAGACTATAAAGATACAGGCGTTTCCCGCT 125
Db |
QY 810 GGAAGCTCCCTGCTGGCGTCTCTCTCCGACCTGCCGCTTACCGGATACCTGTCGCC 869
Db |
QY 126 GGAAGCTCCCTGCTGGCGTCTCTCTCCGACCTGCCGCTTACCGGATACCTGTCGCC 185
Db |
QY 870 TTTCTCCCTTCCGGAAGCGTGGCGCTTCTCAATGCTACGCTAGGTATCTCAGTTGC 929
Db |
QY 186 TTTCTCCCTTCCGGAAGCGTGGCGCTTCTCAATGCTACGCTAGGTATCTCAGTTGC 245
Db |
QY 930 GTGTAGGTCGTTCCGCTCCCAAGCTGGGCTGTGTGACGAAACCCCGCTTACCGGACCG 989
Db |
QY 246 GTGTAGGTCGTTCCGCTCCCAAGCTGGGCTGTGTGACGAAACCCCGCTTACCGGACCG 305
Db |
QY 990 TGCCTTATCCGGTAACATATCGTCTTGAAGTCCAAACCCGGTAAGACACGACTTATCGCA 1049
Db |
QY 306 TGCCTTATCCGGTAACATATCGTCTTGAAGTCCAAACCCGGTAAGACACGACTTATCGCA 365
Db |

QY 1050 CTGCGACAGCCACTGCTAAACAGAGTTAGCAGAGCGAGGATGATGAGCGGTGCTACAGAG 1109
Db 366 CTGCGACAGCCACTGCTAAACAGAGTTAGCAGAGCGAGGATGATGAGCGGTGCTACAGAG 425
QY 1110 TTCTTGAAGTGGTGGCTAACTACGGCTACACTAGAGGACAGTATTGGTATCTGCGCT 1169
Db 426 TTCTTGAAGTGGTGGCTAACTACGGCTACACTAGAGGACAGTATTGGTATCTGCGCT 485
QY 1170 CTGCTGAAGCCAGTTACTCTCGGAAAAAGAGTTGCTAGCTCTTGATCCGCGCAACAAACC 1229
Db 486 CTGCTGAAGCCAGTTACTCTCGGAAAAAGAGTTGCTAGCTCTTGATCCGCGCAACAAACC 545
QY 1230 ACCGCTGGTAGCGGTGTTTTTTTGTGCAAGCAGCAGATTAAGCGCAGAAAAAGGA 1289
Db 546 ACCGCTGGTAGCGGTGTTTTTTTGTGCAAGCAGCAGATTAAGCGCAGAAAAAGGA 605
QY 1290 TCT 1292
Db 606 TCT 608
RESULT 12
PCT-US99-06742-7
; Sequence 7, Application PC/TUS9906742
; GENERAL INFORMATION:
; APPLICANT: Ruffner, Duane E.
; APPLICANT: Pierce, Michael L.
; APPLICANT: Chen, Zhidong
; TITLE OF INVENTION: Directed Antisense Libraries
; FILE REFERENCE: T6678.PCT
; CURRENT APPLICATION NUMBER: PCT/US99/06742
; CURRENT FILING DATE: 1999-03-28
; EARLIER APPLICATION NUMBER: 60/079,792
; EARLIER FILING DATE: 1998-03-28
; NUMBER OF SEQ ID NOS: 34
; SOFTWARE: Word97
; SEQ ID NO 7
; LENGTH: 2077
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: PASLib
PCT-US99-06742-7
Query Match 39.08; Score 603; DB 1; Length 2077;
Best Local Similarity 100.0%; Pred. No. 1e-166;
Matches 603; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 690 GCCGCGTTGCTGGCGTTTTCATAGGCTCCGCCCCCTGACGAGCATCACAAAAATCGA 749
Db 1437 gccgctgtgcgctgttttccatagctccgcccccctgacgagcatcacaataatcga 1496
QY 750 CGCTCAAGTCAGAGGTGGGAAACCCGACAGGACTATAAAGATACACAGCGCTTTCCCGCT 809
Db 1497 cgctcaagtcagaggtgggaaacccgacaggaactataaagataccaggcggtttcccccct 1556
QY 810 GGAAGCTCCCTCGTGGCTCTCTGTTCCGACCTGCGGCTTACCGGATACCTGTCGCC 869
Db 1557 ggaagctccctcgtgcgtctctctgttccgacctgcgcttacccgatacactgctccgcc 1616
QY 870 TTCTTCCCTTCGGGAAGCGTGGCGCTTCTCAATGCTCACGCTAGTATCTCAGTTGCG 929
Db 1617 ttcttcccttcgggaagcgtggcgctttctcaatgctcaacgctcagctgtaggtactcagttcg 1676
QY 930 GTGAGTGTGCTCTCAAGCTGGGCTGTGTGCAAGAACCCCGTTACGCCCGACCGC 989
Db 1677 gtgaggtcgttcgctccaagctggcgctgtgtgcaagaaccccggttccagccgacgcg 1736
QY 990 TGGCGCTTATCCGCTAACTATCTGCTTGTGAGTCCAAACCGGTAAGACAGACTTATCGCCA 1049
Db 1737 tgcgcttataccgctaaactatctgttgcgttcgagtcgaacccggtaagacacgacttatacgcga 1796

QY 1050 CTGCGACAGCCACTGCTAAACAGAGTTAGCAGAGCGAGGATGATGAGCGGTGCTACAGAG 1109
Db 1797 ctggcagcagccactggttaacagagttagcagagcgaggtatgtagcggtgctcacagag 1856
QY 1110 TTCTTGAAGTGGTGGCTAACTACGGCTACACTAGAGGACAGTATTGGTATCTGCGCT 1169
Db 1857 ttcttgaagtgggtgcctaactacggtacactagaaggacagtatgtggtatctcgct 1916
QY 1170 CTGCTGAAGCCAGTTACTCTCGGAAAAAGAGTTGCTAGCTCTTGATCCGCGCAACAAACC 1229
Db 1917 ctgctgaagccagttactcttcggaaaaagagttggtagctcttgatccgcgaacaaacc 1976
QY 1230 ACCGCTGGTAGCGGTGTTTTTTTGTGCAAGCAGCAGATTAAGCGCAGAAAAAGGA 1289
Db 1977 accgctgtagcggtgtttttttgttgcgaagcagcagattacgcgcagaaaaaagga 2036
QY 1290 TCT 1292
Db 2037 tct 2039
RESULT 13
US-09-496-445-5
; Sequence 5, Application US/09496445
; GENERAL INFORMATION:
; APPLICANT: McNeish, John D.
; APPLICANT: Ahljinian, Michael K
; TITLE OF INVENTION: Transgenic Animals Expressing Human p25
; FILE REFERENCE: PC10142A
; CURRENT APPLICATION NUMBER: US/09/496,445
; CURRENT FILING DATE: 2000-02-02
; NUMBER OF SEQ ID NOS: 7
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 5
; LENGTH: 2462
; TYPE: DNA
; ORGANISM: Escherichia coli
US-09-496-445-5
Query Match 39.08; Score 603; DB 18; Length 2462;
Best Local Similarity 100.0%; Pred. No. 1.1e-166;
Matches 603; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 690 GCCGCGTTGCTGGCGTTTTCATAGGCTCCGCCCCCTGACGAGCATCACAAAAATCGA 749
Db 359 gccgctgtgcgctgttttccatagctccgcccccctgacgagcatcacaataatcga 418
QY 750 CGCTCAAGTCAGAGGTGGGAAACCCGACAGGACTATAAAGATACACAGCGCTTTCCCGCT 809
Db 419 cgctcaagtcagaggtgggaaacccgacaggaactataaagataccaggcggtttcccccct 478
QY 810 GGAAGCTCCCTCGTGGCTCTCTGTTCCGACCTGCGGCTTACCGGATACCTGTCGCC 869
Db 479 ggaagctccctcgtgcgtctctctgttccgacctgcgcttacccgatacactgctccgcc 538
QY 870 TTCTTCCCTTCGGGAAGCGTGGCGCTTCTCAATGCTCACGCTAGTATCTCAGTTGCG 929
Db 539 ttcttcccttcgggaagcgtggcgctttctcaatgctcaacgctgtaggtatacctcagtttcg 598
QY 930 GTGAGTGTGCTCTCAAGCTGGGCTGTGTGCAAGAACCCCGTTACGCCCGACCGC 989
Db 599 gtgaggtcgttcgctccaagctggcgctgtgtgcaagaaccccggttccagccgacgcg 658
QY 990 TGGCGCTTATCCGCTAACTATCTGCTTGTGAGTCCAAACCGGTAAGACAGACTTATCGCCA 1049
Db 659 tgcgcttataccgctaaactatctgttgcgttcgaaccccggttaagacacgacttatacgcga 718
QY 1050 CTGCGACAGCCACTGCTAAACAGAGTTAGCAGAGCGAGGATGATGAGCGGTGCTACAGAG 1109
Db 719 ctggcagcagccactggttaacagagttagcagagcgaggtatgtagcggtgctcacagag 778
QY 1110 TTCTTGAAGTGGTGGCTAACTACGGCTACACTAGAGGACAGTATTGGTATCTGCGCT 1169

Db 779 ttcttgaagtgtggtcctaacctacggtacactagaagcagatttggatctgcgt 838
 QY 1170 CTGCTGAAGCCAGTTACCTTCGGAAGAGAGTTGTTAGCTCTTGTATCGCGCAAAAC 1229
 Db 839 ctgctgaagccagttaccttcggaagagcttggtgctcttgatccgcgaacaaacc 898
 QY 1230 ACCGCTGTAGCGTGTGTTTTTTTGTGCAAGCAGCAGATTACGCGCAGAAAAAGGA 1289
 Db 899 accgctgtgagcgtgtgtttttttgttttgaagcagcagattacgcgcagaaaaagga 958
 QY 1290 TCT 1292
 Db 959 tct 961
 RESULT 14
 US-60-118-478-5
 ; Sequence 5, Application US/60118478
 ; GENERAL INFORMATION:
 ; APPLICANT: McNeish, John D.
 ; APPLICANT: Ahlman, Michael K
 ; TITLE OF INVENTION: Transgenic Animals Expressing Human p25
 ; FILE REFERENCE: PC10142
 ; CURRENT APPLICATION NUMBER: US/60/118,478
 ; CURRENT FILING DATE: 1999-02-03
 ; NUMBER OF SEQ ID NOS: 7
 ; SOFTWARE: PatentIn Ver. 2.0 - beta
 ; SEQ ID NO 5
 ; LENGTH: 2462
 ; TYPE: DNA
 ; ORGANISM: Escherichia coli
 US-60-118-478-5

Query Match 39.0%; Score 603; DB 44; Length 2462;
 Best Local Similarity 100.0%; Pred. No. 1.le-166;
 Matches 603; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 690 GCCGCGTGTGCGGTTTTCATAGCTCCGCCCTGACGAGCATCACAAAAATCGA 749
 Db 359 gccgctgtgctgtgttttccatagctccgccctgacgagcatcaaaaaatcga 418
 QY 750 CGCTCAAGTCAGAGTGGCGAACCACGACGACTATAAGATACACAGGCTTCCCT 809
 Db 419 cgtcaagtcagagtggtggaacccgacagagactataagataccaggcgtttccct 478
 QY 810 GGAAGCTCCCTGCTGCGCTCTCTGTTCCGACCTCGCGCTTACCGGATACCTGTCGCC 869
 Db 479 ggaagctccctgtgcgtctctctgttccgacctgctccggtacccgtacccgtccgc 538
 QY 870 TTCTCCCTTCGGGAAGCGTGGCGTTTCTCAATGCTACGCTGTAGGTATCTCAGTTCG 929
 Db 539 ttctcccttcggaagcgtggcgcttctcaatgtcagctgtaggtatctcagttcg 598
 QY 930 GTGTAGTGTGCTGCCGCTCCAGCTGGGCTGTGCGAGAACCCCGCTTACGCGACCGC 989
 Db 599 ggtgagtgctgcgtcccaagctgggctgtgtgcaagaccccgcttccagccgacgc 658
 QY 990 TCGCCTTATCCGGTAACCTATCTGTTAGTCCAAACCGGTAAGACACGACTATCGCCA 1049
 Db 659 tgcgcttaccggttaactatctgttctgagtcaccccggttaagacagactatcgcca 718
 QY 1050 CTGGCAGCAGCCACTGGTAACAGGATTAGCAGAGGAGGTATGAGGCGGTCTACAGAG 1109
 Db 719 ctggcagcagccactggtaacaggtattagcagagcaggtatgagcgtgctacagag 778
 QY 1110 TTCTTGAAGTGTGCGCTTAACCTACGCTACACTAGAAGCAGAGTATTGGTATCGCGCT 1169
 Db 779 ttcttgaagtgtggtcctaacctacggtacactagaagacagatttggatctgcgt 838
 QY 1170 CTGCTGAAGCCAGTTACCTTCGGAAGAGAGTTGTTAGCTCTTGTATCGCGCAAAAC 1229

Db 839 ctgctgaagccagttaccttcggaagaggttggtagctcttgatccgcgaacaaacc 898
 QY 1230 ACCGCTGTAGCGTGTGTTTTTTTGTGCAAGCAGCAGATTACGCGCAGAAAAAGGA 1289
 Db 899 accgctgtgagcgtgtgtttttgttttgaagcagcagattacgcgcagaaaaagga 958
 QY 1290 TCT 1292
 Db 959 tct 961
 RESULT 15
 US-07-919-535C-26
 ; Sequence 26, Application US/07919535C
 ; GENERAL INFORMATION:
 ; APPLICANT: Haas, Werner
 ; APPLICANT: Hunziker, Willi
 ; TITLE OF INVENTION: SOLUBLE KIT LIGANDS
 ; NUMBER OF SEQUENCES: 68
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: Hoffmann-La Roche Inc.
 ; STREET: 340 Kingsland Street
 ; CITY: Nutley
 ; STATE: New Jersey
 ; COUNTRY: United States
 ; ZIP: 07110-1199
 ; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: Floppy disk
 ; COMPUTER: IBM PC compatible
 ; OPERATING SYSTEM: DOS 4.0
 ; SOFTWARE: PatentIn Release #1.0, Version #1.25
 ; CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: US/07/919,535C
 ; FILING DATE: 19930723
 ; CLASSIFICATION: 435
 ; PRIOR APPLICATION DATA:
 ; APPLICATION NUMBER: EP 91810609.7
 ; FILING DATE: 30-JUL-1991
 ; ATTORNEY/AGENT INFORMATION:
 ; NAME: Kass, Alan P.
 ; REGISTRATION NUMBER: 32142
 ; REFERENCE/DOCKET NUMBER: RAN 4105/142
 ; TELEPHONE: (201) 235-4205
 ; TELEFAX: (201) 235-3500
 ; INFORMATION FOR SEQ ID NO: 26:
 ; SEQUENCE CHARACTERISTICS:
 ; LENGTH: 2577 base pairs
 ; TYPE: NUCLEIC ACID
 ; STRANDEDNESS: double
 ; TOPOLOGY: linear
 ; MOLECULE TYPE: DNA (genomic)
 ; ORIGINAL SOURCE:
 ; ORGANISM: Escherichia coli
 ; FEATURE:
 ; NAME/KEY: misc_feature
 ; LOCATION: 121..126
 ; OTHER INFORMATION: /note= "Recognition site for
 ; OTHER INFORMATION: restriction enzyme NcoI"
 ; FEATURE:
 ; NAME/KEY: misc_feature
 ; LOCATION: 141..146
 ; OTHER INFORMATION: /note= "Recognition site for
 ; OTHER INFORMATION: restriction enzyme HindIII"
 ; FEATURE:
 ; NAME/KEY: misc_feature
 ; LOCATION: 1..2577
 ; OTHER INFORMATION: /note= "Complete nucleotide
 ; OTHER INFORMATION: sequence of E. coli
 ; OTHER INFORMATION: expression plasmid
 ; OTHER INFORMATION: POC 56/RBS II, NcoI"
 US-07-919-535C-26

Query Match		39.0%;	Score 603;	DB 3;	Length 2577;				
Best Local Similarity		100.0%;	Pred. No. 1.1e-166;						
Matches 603;		Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;				
Qy	690	GCCGCGTTGCTGGCGTTTTTCATAGGCTCGGCCCTGACGAGCATCAAAAAATCGA	749						
Db	736	GCCGCGTTGCTGGCGTTTTTCATAGGCTCGGCCCTGACGAGCATCAAAAAATCGA	795						
Qy	750	CGTCTAAGTCAGAGTGGCGAAACCCGACAGGACTATAAAGATACACAGGCGTTTCCGCCCT	809						
Db	796	CGTCTAAGTCAGAGTGGCGAAACCCGACAGGACTATAAAGATACACAGGCGTTTCCGCCCT	855						
Qy	810	GGAAGCTCCCTCGTGGCGCTCTCTGTTTCGACCTCGCGCTTACCGGATACCTGTCCGCC	869						
Db	856	GGAAGCTCCCTCGTGGCGCTCTCTGTTTCGACCTCGCGCTTACCGGATACCTGTCCGCC	915						
Qy	870	TTTCTCCCTTCGGGAAGCGTGGCGCTTTCATAATGCTACGCTGTAGGTATCTCAGTTTCG	929						
Db	916	TTTCTCCCTTCGGGAAGCGTGGCGCTTTCATAATGCTACGCTGTAGGTATCTCAGTTTCG	975						
Qy	930	GTGTAGGTGCTTCGCTCCAAGCTGGGCTGTGTCGACCAACCCCGCTTCAGCCGACGCG	989						
Db	976	GTGTAGGTGCTTCGCTCCAAGCTGGGCTGTGTCGACCAACCCCGCTTCAGCCGACGCG	1035						
Qy	990	TGCGCCTTATCCGGTAAGTCTGCTTGTAGTCCAAACCCGGTAAAGACACGACTTATCGCCA	1049						
Db	1036	TGCGCCTTATCCGGTAAGTCTGCTTGTAGTCCAAACCCGGTAAAGACACGACTTATCGCCA	1095						
Qy	1050	CTGGCAGCAGCCACTGGTAAACAGGATTAGCAGACGAGGATGTAGCGGTGTGTACAGAG	1109						
Db	1096	CTGGCAGCAGCCACTGGTAAACAGGATTAGCAGACGAGGATGTAGCGGTGTGTACAGAG	1155						
Qy	1110	TTCTTGAAGTGGTGGCCCTAACTACGGCTACACTAGAAAGGACAGTATTGGTATCTGCGCT	1169						
Db	1156	TTCTTGAAGTGGTGGCCCTAACTACGGCTACACTAGAAAGGACAGTATTGGTATCTGCGCT	1215						
Qy	1170	CTGCTGAAGCAGTTACCTTCGGAAGAGTTGGTACTTGTATCCGGCAACAAACC	1229						
Db	1216	CTGCTGAAGCAGTTACCTTCGGAAGAGTTGGTACTTGTATCCGGCAACAAACC	1275						
Qy	1230	ACCGCTGGTAGCGGTGTTTTTTTGTGCAAGCAGAGATTACCGCCAGAAAAAAGGA	1289						
Db	1276	ACCGCTGGTAGCGGTGTTTTTTTGTGCAAGCAGAGATTACCGCCAGAAAAAAGGA	1335						
Qy	1290	TCT 1292							
Db	1336	TCT 1338							

Search completed: January 17, 2002, 15:14:01
Job time: 22753 sec

THIS PAGE BLANK (USPTO)

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: January 17, 2002, 08:54:22 ; Search time 6501.33 Seconds

(without alignments)
2556.972 Million cell updates/sec

Title: US-09-242-202a-27

Perfect score: 1547

Sequence: 1 GGTACCGCCACCATGGCGC.....CTGCACAGCGCTCTCCACACA 1547

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 11351937 seqs, 5372889281 residues

Total number of hits satisfying chosen parameters: 22703874

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

EST:*
1: em_estfun:*
2: em_esthum:*
3: em_estin:*
4: em_estom:*
5: em_estpl:*
6: em_estba:*
7: em_estro:*
8: em_estov:*
9: em_htc:*
10: gb_estl1:*
11: gb_est2:*
12: gb_htc:*
13: gb_gss:*
14: em_gss_fun:*
15: em_gss_hum:*
16: em_gss_inv:*
17: em_gss_pln:*
18: em_gss_pro:*
19: em_gss_tod:*
20: em_gss_vrt:*
21: em_gss_other:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	597.2	38.6	1004	10	AJ281480
2	596.6	38.6	800	10	AJ281449
3	592.2	38.3	1070	10	AJ281552
C 4	585.4	37.8	756	10	BE749097
C 5	583.8	37.7	959	10	BE749147
C 6	580.4	37.5	759	10	BE749118
C 7	579	37.4	840	10	BE749178
C 8	576.2	37.2	784	13	AQ876119
C 9	573.6	37.1	795	13	AQ876011
C 10	565.2	36.5	617	10	AJ281661
11	555.2	35.9	571	10	AL044178
12	554	35.8	579	10	AJ281320

C 13	554	35.8	794	13	AQ875988
14	553.2	35.8	615	10	AV735664
15	550.2	35.6	954	10	AL044364
16	549.6	35.5	1067	10	AU081137
C 17	549.4	35.5	706	10	BE268532
18	549.2	35.5	579	10	AL043613
C 19	548.6	35.5	1045	11	BF664532
20	548.2	35.4	629	10	AL593919
21	547.2	35.4	568	10	AJ281376
22	546.2	35.3	700	10	AJ281616
C 23	544.8	35.2	982	10	BE749172
24	544.6	35.2	795	13	AQ875935
25	543.2	35.1	548	10	AJ281654
26	543.2	35.1	718	13	AG010489
27	536.2	34.7	616	10	AV735756
28	532.2	34.4	615	10	AL044413
C 29	529.2	34.2	857	13	AQ875006
C 30	528.6	34.2	747	10	BE749123
C 31	524.4	33.9	890	11	BG685100
C 32	523	33.8	884	13	AQ491919
C 33	521.2	33.7	526	10	AL043840
C 34	521.2	33.7	852	13	AQ491918
C 35	518.2	33.5	1089	10	AU081124
C 36	512.4	33.1	691	10	BE749090
C 37	509.2	32.9	798	13	AQ876257
38	507	32.8	636	10	AJ281699
39	506.2	32.7	628	11	BF381364
C 40	505.4	32.7	902	10	BE749089
41	502.4	32.5	630	10	AL042334
42	500.4	32.3	608	13	AG000621
43	500.2	32.3	527	10	AL043585
C 44	499	32.3	799	13	AQ876001
C 45	497.6	32.2	950	10	BE737278

ALIGNMENTS

RESULT	1	AJ281480	1004 bp	mrna	EST	30-JUN-2000
AJ281480	LOCUS	4A3A-P4G8-F	Anopheles gambiae	immune competent	4A3A Anopheles	
DEFINITION		gambiae cDNA clone 4A3A-P4G8,	mrna	sequence.		
ACCESSION		AJ281480				
VERSION		AJ281480.1	GI:6929360			
KEYWORDS		EST.				
SOURCE		African malaria mosquito.				
ORGANISM		Anopheles gambiae				
REFERENCE		1 (bases 1 to 1004)				
AUTHORS		Dimopoulos G., Casavant, T.L., Chang, S., Scheetz, T., Roberts, C., Donohue, M., Schultz, J., Benes, V., Bork, P., Ansorge, W., Soares, M.B. and Kafatos, F.C.				
TITLE		Anopheles gambiae pilot gene discovery project: identification of mosquito innate immunity genes from expressed sequence tags generated from immune-competent cell lines				
JOURNAL		Proc. Natl. Acad. Sci. U.S.A.	97 (12),	6619-6624	(2000)	
MEDLINE		20300950				
COMMENT		Contact: Dimopoulos G				
FEATURES		source				
		1. .1004				
		/organism="Anopheles gambiae"				
		/strain="4A r/r"				
		/db_xref="taxon:7165"				
		/clone="4A3A-P4G8"				
		/clone_lib="Anopheles gambiae immune competent 4A3A"				
		/cell_line="immune competent 4A3A"				
		/lab_host="E. coli DH10B"				

/note="vector: p7T3D-Pac (Pharmacia) with a modified
polylinker; Site_1: EcoRI; Site_2: NotI; sequenced from
forward priming site which reads from the 3' end of the
cDNA. The 4A3A is a directionally cloned and normalized
cDNA library that was constructed from the 4A3A cell line
oligo-T primed cDNA according to: Bonaldo, Lennon & Soares
(1996) : Normalization and Subtraction: Two approaches To
Facilitate Gene Discovery, Genome Research 6, 791-806."

BASE COUNT 252 a 262 c 244 g 244 t 2 others
ORIGIN

Query Match 38.6%; Score 597.2; DB 10; Length 1004;
Best Local Similarity 99.3%; Pred. No. 1e-160;
Matches 599; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

Qy 690 GCCGGTGTGCGCTTTTCATAGGCTCGCCCGCTGACGAGCATCACAAAATCGA 749
Db 24 GCCGGTGTGCGCTTTTCATAGGCTCGCCCGCTGACGAGCATCACAAAATCGA 83
Qy 750 CGCTCAAGTCAGAGTGGCGAAACCCGACAGGACTATAAGATACAGCGGTTCGCCCT 809
Db 84 CGCTCAAGTCAGAGTGGCGAAACCCGACAGGACTATAAGATACAGCGGTTCGCCCT 143
Qy 810 GGAAGCTCCCTGCGCTCTCTGTCGACCGCTGCGCTTACGGATACGTGTCGCC 869
Db 144 GGAAGCTCCCTGCGCTCTCTGTCGACCGCTGCGCTTACGGATACGTGTCGCC 203
Qy 870 TTTCTCCCTCGGGAAGGTGGCGCTTCTCAATGCTCACGCTGTAGTATCTCAGTTTCG 929
Db 204 TTTCTCCCTCGGGAAGGTGGCGCTTCTCATAGCTCACGCTGTAGTATCTCAGTTTCG 263
Qy 930 GTGTAGGTGCTTCCGCTCCTCAAGCTGGGCTGTGTGCACGAAACCCCGCTTACGCCGACCGC 989
Db 264 GTGTAGGTGCTTCCGCTCCTCAAGCTGGGCTGTGTGCACGAAACCCCGCTTACGCCGACCGC 323
Qy 990 TGGCGCTTATCCGCTAATCTGCTTCAGTCCACCCGCTTAAGACAGACTTATCGCCA 1049
Db 324 TGGCGCTTATCCGCTAATCTGCTTCAGTCCACCCGCTTAAGACAGACTTATCGCCA 383
Qy 1050 CTGGCAGCAGCACTGGTAACAGGATTAGCAGAGGAGTATGTAGCGGTGCTACAGAG 1109
Db 384 CTGGCAGCAGCACTGGTAACAGGATTAGCAGAGGAGTATGTAGCGGTGCTACAGAG 443
Qy 1110 TTTCTGAAGTGGTGGCTTAATCTAGGCTTACACTAGAGGAGCAGTATTTGGTATCTGCGCT 1169
Db 444 TTTCTGAAGTGGTGGCTTAATCTAGGCTTACACTAGAGGAGCAGTATTTGGTATCTGCGCT 503
Qy 1170 CTGCTGAAGCCAGTTACCTTCGAAAGAGTGTGTAGTCTTGTATCCGCAACAAACC 1229
Db 504 CTGCTGAAGCCAGTTACCTTCGAAAGAGTGTGTAGTCTTGTATCCGCAACAAACC 563
Qy 1230 ACCGCTGTAGCGGTGGTGTGTTTTGTTTTCGAAAGCAGCAGATTTACGCGCAGAAAAAGGA 1289
Db 564 ACCGCTGTAGCGGTGGTGTGTTTTGTTTTCGAAAGCAGCAGATTTACGCGCAGAAAAAGGA 623
Qy 1290 TCT 1292
Db 624 TCT 626

RESULT 2
AJ281449 800 bp mRNA EST 30-JUN-2000
LOCUS 4A3A-P4D5-F Anopheles gambiae immune competent 4A3A Anopheles
DEFINITION gambiae cDNA clone 4A3A-P4D5, mRNA sequence.
ACCESSION AJ281449
VERSION AJ281449.1 GI:5929329
KEYWORDS EST.
SOURCE African malaria mosquito.
ORGANISM Anopheles gambiae
Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
Pterygota; Neoptera; Endopterygota; Diptera; Nematocera; Culicoidea

REFERENCE : Anopheles.
AUTHORS 1 (bases 1 to 800)
Donohue, G., Casavant, T.L., Chang, S., Scheetz, T., Roberts, C.,
and Kafatos, F.C.
TITLE Anopheles gambiae pilot gene discovery project: identification of
mosquito innate immunity genes from expressed sequence tags
generated from immune-competent cell lines
JOURNAL Proc. Natl. Acad. Sci. U.S.A. 97 (12), 6619-6624 (2000)
MEDLINE 20300950
COMMENT Contact: Dimopoulos G
Fotis C. Kafatos laboratory
European Molecular Biology Laboratory
Meyerhofstrasse 1, 69117 Heidelberg, Germany.
Location/Qualifiers
FEATURES
source

1. 800
/organism="Anopheles gambiae"
/strain="4A r/r"
/db_xref="taxon:7165"
/clone="4A3A-P4D5"
/cell_line="Immune competent 4A3A"
/lab_host="E. coli DH10b"
/note="Vector: p7T3D-Pac (Pharmacia) with a modified
polylinker; Site_1: EcoRI; Site_2: NotI; sequenced from
forward priming site which reads from the 3' end of the
cDNA. The 4A3A is a directionally cloned and normalized
cDNA library that was constructed from the 4A3A cell line
oligo-T primed cDNA according to: Bonaldo, Lennon & Soares
(1996) : Normalization and Subtraction: Two approaches To
Facilitate Gene Discovery, Genome Research 6, 791-806."
BASE COUNT 203 a 206 c 198 g 193 t
ORIGIN

Query Match 38.6%; Score 596.6; DB 10; Length 800;
Best Local Similarity 99.3%; Pred. No. 1.4e-160;
Matches 599; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

Qy 690 GCCGGTGTGCGCTTTTCATAGGCTCGCCCGCTGACGAGCATCACAAAATCGA 749
Db 51 GCCGGTGTGCGCTTTTCATAGGCTCGCCCGCTGACGAGCATCACAAAATAGA 110
Qy 750 CGCTCAAGTCAGAGTGGCGAAACCCGACAGGACTATAAGATACAGCGGTTCGCCCT 809
Db 111 CGCTCAAGTCAGAGTGGCGAAACCCGACAGGACTATAAGATACAGCGGTTCGCCCT 170
Qy 810 GGAAGCTCCCTGCGCTCTCTTCCGACCCCTGCGCTTACCGGATACCTGTCCGCC 869
Db 171 GGAAGCTCCCTGCGCTCTCTTCCGACCCCTGCGCTTACCGGATACCTGTCCGCC 230
Qy 870 TTTCTCCCTTCGGGAAGCGTGGCGCTTTCTCAATGCTCACGCTGTAGTATCTCAGTTTCG 929
Db 231 TTTCTCCCTTCGGGAAGCGTGGCGCTTTCTCATAGCTCACGCTGTAGTATCTCAGTTTCG 290
Qy 930 GTGTAGGTGCTGCGTCCAAAGCTGGGCTGTGTGCAGAACCCCGCTTACGCCGACCGC 989
Db 291 GTGTAGGTGCTGCGTCCAAAGCTGGGCTGTGTGCAGAACCCCGCTTACGCCGACCGC 350
Qy 990 TGGCGCTTATCCGCTAATCTGCTTCAGTCCACCCGCTTAAGACAGACTTATCGCCA 1049
Db 351 TGGCGCTTATCCGCTAATCTGCTTCAGTCCACCCGCTTAAGACAGACTTATCGCCA 410
Qy 1050 CTGGCAGCAGCACTGGTAACAGGATTAGCAGAGGAGTATGTAGCGGTGCTACAGAG 1109
Db 411 CTGGCAGCAGCACTGGTAACAGGATTAGCAGAGGAGTATGTAGCGGTGCTACAGAG 470
Qy 1110 TTTCTGAAGTGGTGGCTTAATCTAGGCTTACACTAGAGGAGCAGTATTTGGTATCTGCGCT 1169
Db 471 TTTCTGAAGTGGTGGCTTAATCTAGGCTTACACTAGAGGAGCAGTATTTGGTATCTGCGCT 530
Qy 1170 CTGCTGAAGCCAGTTACCTTCGAAAGAGTGTGTAGTCTTGTATCCGCAACAAACC 1229

```

Db 531 CTGCTGAAGCCAGTTACTCTCGGAAAAAGAGTTGGTAGCTCTTGTATCCGGCAACAAACC 590
Qy 1230 ACCCTGGTAGCGGTGTTTTTTTGGTTTGAAGCAGCAGATTACGCCGACAGAAAAAGGA 1289
Db 591 ACCCTGGTAGCGGTGTTTTTTTGGTTTGAAGCAGCAGATTACGCCGACAGAAAAAGGA 650
Qy 1290 TCT 1292
Db 651 TCT 653

RESULT 3
AJ281552
LOCUS
DEFINITION
4A3A-P6F11-F Anopheles gambiae immune competent 4A3A Anopheles
gambiae cDNA clone 4A3A-P6F11, mRNA sequence.
ACCESSION
AJ281552
VERSION
AJ281552.1 GI:6929432
KEYWORDS
EST.
SOURCE
African malaria mosquito.
ORGANISM
Anopheles gambiae
Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
Pterygota; Neoptera; Endopterygota; Diptera; Nematocera; Culicoidea
; Anopheles.
REFERENCE
1 (bases 1 to 1070)
Donohue M.G., Casavant, T.L., Chang, S., Scheetz, T., Roberts, C.,
Donohue, M., Schultz, J., Benes, V., Bork, P., Ansorge, W., Soares, M.B.
and Kafatos, F.C.
Anopheles gambiae pilot gene discovery project: identification of
mosquito innate immunity genes from expressed sequence tags
generated from immune-competent cell lines
Proc. Natl. Acad. Sci. U.S.A. 97 (12), 6619-6624 (2000)
20300950
Contact: Dimopoulos G
Fotis C. Kafatos Laboratory
European Molecular Biology Laboratory
Meyerhofstrasse 1, 69117 Heidelberg, Germany.
FEATURES
Location/Qualifiers
1..1070
/organism="Anopheles gambiae"
/strain="4A r"
/db_xref="taxon:7165"
/clone="4A3A-P6F11"
/cell_line="immune competent 4A3A"
/lab_host="E. coli DH10B"
/notes="Vector: pF73D-Pac (Pharmacia) with a modified
polylinker; Site_1: EcoRI; Site_2: NotI; sequenced from
forward priming site which reads from the 3' end of the
cDNA. The 4A3A is a directionally cloned and normalized
cDNA library that was constructed from the 4A3A cell line
oligo-T primed cDNA according to: Bonaldo, Lennon & Soares
(1996) : Normalization and Subtraction: Two approaches to
Facilitate Gene Discovery, Genome Research 6, 791-806."
BASE COUNT 263 a 283 c 255 g 269 t
ORIGIN

Query Match 38.3%; Score 592.2; DB 10; Length 1070;
Best Local Similarity 99.5%; Pred. No. 2.9e-159;
Matches 594; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy 696 TTGCTGGGTTTTCATAGCTCCGCCCTCAGCAGCATCACAAAATCGAGCTCA 755
Db 1 TTGCTGGGTTTTCATAGCTCCGCCCTCAGCAGCATCACAAAATCGAGCTCA 60
Qy 756 AGTCAGAGTGGCAAAACCCGACAGGACTATAAAGATACCAGGGTTTCCCTCGGAAGC 815
Db 61 AGTCAGAGTGGCAAAACCCGACAGGACTATAAAGATACCAGGGTTTCCCTCGGAAGC 120
Qy 816 TCCCTGTGCGCTCTCTGTGTCGACCGCTGCGGTACCGGATACCTGTGCGCTTCTC 875
Db 121 TCCCTGTGCGCTCTCTGTGTCGACCGCTGCGGTACCGGATACCTGTGCGCTTCTC 180

```

```

Qy 876 CCTTCGGGAAGCGTGGCGCTTTTCAATGCTCAGCGTGTAGTATCTCAGTTCGGTGTAG 935
Db 181 CCTTCGGGAAGCGTGGCGCTTTTCAATGCTCAGCGTGTAGTATCTCAGTTCGGTGTAG 240
Qy 936 GTCGTTGCTCCAAAGCTGGGCTGTGTGACGAACCCCGTTACGCGGACCGCTGGGCC 995
Db 241 GTCGTTGCTCCAAAGCTGGGCTGTGTGACGAACCCCGTTACGCGGACCGCTGGGCC 300
Qy 996 TTATCCGGTAACTATCGTCTTGAAGTCCAAACCCGGTAAACACGACTTATCCGCACTGSCA 1055
Db 301 TTATCCGGTAACTATCGTCTTGAAGTCCAAACCCGGTAAACACGACTTATCCGCACTGSCA 360
Qy 1056 GCAGCCACTGTTAAGGATTAGCAGAGCGAGGTATGTAGCGGTGTACAGAGTTCTTG 1115
Db 361 GCAGCCACTGTTAAGGATTAGCAGAGCGAGGTATGTAGCGGTGTACAGAGTTCTTG 420
Qy 1116 AAGTGGTGGCCTAACTACGGCTACACTAGAGAGCAGATTGTTGGTATCTGCGCTGTGCTG 1175
Db 421 AAGTGGTGGCCTAACTACGGCTACACTAGAGAGCAGATTGTTGGTATCTGCGCTGTGCTG 480
Qy 1176 AAGCCAGTTACTCTCGGAAAAGAGTTGGTAGCTCTTCGCGCAACAAACACCGCT 1235
Db 481 AAGCCAGTTACTCTCGGAAAAGAGTTGGTAGCTCTTCGCGCAACAAACACCGCT 540
Qy 1236 GGTAGCGGTGTTTTTTTGTGTTGCAAGCAGCAGATTACCGCGCAACAAAGGATCT 1292
Db 541 GGTAGCGGTGTTTTTTTGTGTTGCAAGCAGCAGATTACCGCGCAACAAAGGATCT 597

RESULT 4
BE749097/c
LOCUS
DEFINITION
601123138F1 NIH_MGC_5 Homo sapiens cDNA clone IMAGE:3348224 5',
mRNA sequence.
ACCESSION
BE749097
VERSION
BE749097.1 GI:10163089
KEYWORDS
EST.
SOURCE
human.
ORGANISM
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE
1 (bases 1 to 756)
NIH-MGC http://mgc.nci.nih.gov/.
AUTHORS
National Institutes of Health, Mammalian Gene Collection (MGC)
TITLE
Unpublished (1999)
JOURNAL
Contact: Robert Strausberg, Ph.D.
COMMENT
Email: cgabbs-r@mail.nih.gov
Tissue Procurement: ATCC
CDNA Library Preparation: Ling Hong/Rubin Laboratory
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at: image.llnl.gov
Plate: LLCMI40 row: k column: 09
High quality sequence start: 7
High quality sequence stop: 739.
FEATURES
Location/Qualifiers
1..756
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:3348224"
/cell_line="NIH_MGC_5"
/tissue_type="carcinoma cell line"
/lab_host="DH10B (phage-resistant)"
/notes="Organ: cervix; Vector: pORF7; Site_1: XhoI; Site_2:
EcoRI; cDNA made by oligo-dT priming. Directionally
cloned into EcoRI/XhoI sites using the following 5'
adaptor: GCGACGAG(G). Size-selected >500bp for average
insert size 1.8kb. Library constructed by Ling Hong in
the laboratory of Gerald M. Rubin (University of
California, Berkeley) using ZAP-cDNA synthesis kit
(Stratagene) and Superscript II RT (Life Technologies)."

```



```

Db 131 TTACGGCAGAAAAAGGATCT 109

RESULT 6
BE749118/c
LOCUS BE749118.1 759 bp mRNA EST 15-SEP-2000
DEFINITION 601123194f1 NIH_MGC_5 Homo sapiens cDNA clone IMAGE:3348238 5',
mRNA sequence.
ACCESSION BE749118
VERSION BE749118.1 GI:10163110
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
AUTHORS 1 (bases 1 to 759)
TITLE NIH-MGC http://mgc.nci.nih.gov/.
JOURNAL National Institutes of Health, Mammalian Gene Collection (MGC)
COMMENT Unpublished (1999)
Contact: Robert Strausberg, Ph.D.
Email: rcgaps@rmail.nih.gov
Tissue Procurement: ATCC
CDNA Library Preparation: Ling Hong/Rubin Laboratory
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at: image.llnl.gov
Plate: LLCM140 row: k column: 23
High quality sequence start: 8
High quality sequence stop: 747.
Location/Qualifiers
1. 759
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:3348238"
/tissue_type="carcinoma cell line"
/lab_host="NIH_MGC_5"
/note="Organ: cervix; Vector: pOTB7; Site_1: XhoI; Site_2:
EcoRI; CDNA made by oligo-dT priming. Directionally
cloned into EcoRI/XhoI sites using the following 5'
adaptor: GGCACGAG(G). Size-selected >500bp for average
insert size 1.8kb. Library constructed by Ling Hong in
the laboratory of Gerald M. Rubin (University of
California, Berkeley) using ZAP-cDNA synthesis kit
(Stratagene) and Superscript II RT (Life Technologies)."
BASE COUNT 191 a 193 c 203 g 172 t
ORIGIN
Query Match 37.5%; Score 580.4; DB 10; Length 759;
Best Local Similarity 98.7%; Pred. No. 6.6e-156;
Matches 596; Conservative 0; Mismatches 6; Indels 2; Gaps 1;

QY 691 CCGGCTTGTGGCTTTTCCATAGGCTCCGCCCTGACGAGCATC--ACAAAAATCG 748
DB 734 CCGGCTTGTGGCTTTTCCATAGGCTCCGCCCTGACGAGCATCTACTAATATCG 675
QY 749 ACGCTCAAGTCAGAGTGGCGAAACCGACAGAGACTATAAGATACACGAGGCTTTCCGCC 808
DB 674 ACGCTCAAGTCAGAGTGGCGAAACCGACAGAGACTATAAGATACACGAGGCTTTCCGCC 615
QY 809 TGAAGCTCCCTCGTGGCTCTCTGTTTCCGACCCCTGACGATACCTGTCGCCG 868
DB 614 TGAAGCTCCCTCGTGGCTCTCTGTTTCCGACCCCTGACGATACCTGTCGCCG 555
QY 869 CTTTCTCCCTTCCGAAAGCGTGGCGCTTCTCAATGCTACGCTGTAGGTATCTCAGTTC 928
DB 554 CTTTCTCCCTTCCGAAAGCGTGGCGCTTCTCAATGCTACGCTGTAGGTATCTCAGTTC 495
QY 929 GGTGTAGTCTGCTGCTCCAGCTGGGCTGTGTGTCAGCAACCCCGCTTACGCGGACCG 988
DB 494 GGTGTAGTCTGCTGCTCCAGCTGGGCTGTGTGTCAGCAACCCCGCTTACGCGGACCG 435

```

```

QY 989 CTGCGCCTTATCCGGTAACATATCGTCTTGTAGTCCAAACCCGTAAGACACGACTTATCGCC 1048
DB 434 CTGCGCCTTATCCGGTAACATATCGTCTTGTAGTCCAAACCCGTAAGACACGACTTATCGCC 375
QY 1049 ACTGGCAGCAGCCACTGGTAAACAGGATTAGCAGAGCGAGGTATGTAGCGGTGCTACAGA 1108
DB 374 ACTGGCAGCAGCCACTGGTAAACAGGATTAGCAGAGCGAGGTATGTAGCGGTGCTACAGA 315
QY 1109 GTTCTTGAAGTGGTGGCCCTAACTACGGCTACACTAGAAGGACAGTATTGTTGTAATCTGCGC 1168
DB 314 GTTCTTGAAGTGGTGGCCCTAACTACGGCTACACTAGAAGGACAGTATTGTTGTAATCTGCGC 255
QY 1169 TCTGCTGAAGCAGGTTACTCTCGGAAAAAGAGTTGGTAGCTCTTGTATCGCGCAAAACAAC 1228
DB 254 TCTGCTGAAGCAGGTTACTCTCGGAAAAAGAGTTGGTAGCTCTTGTATCGCGCAAAACAAC 195
QY 1229 CACCGCTGGTAGCGGTGGTGTGTTTGTGCAAGCAGAGATTACGCGCAGAAAAAAGG 1288
DB 194 CACCGCTGGTAGCGGTGGTGTGTTTGTGCAAGCAGAGATTACGCGCAGAAAAAAGG 135
QY 1289 ATCT 1292
DB 134 ATCT 131

RESULT 7
BE749178/c
LOCUS BE749178.1 840 bp mRNA EST 15-SEP-2000
DEFINITION 601123444f1 NIH_MGC_5 Homo sapiens cDNA clone IMAGE:3348154 5',
mRNA sequence.
ACCESSION BE749178
VERSION BE749178.1 GI:10163170
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
AUTHORS 1 (bases 1 to 840)
TITLE NIH-MGC http://mgc.nci.nih.gov/.
JOURNAL National Institutes of Health, Mammalian Gene Collection (MGC)
COMMENT Unpublished (1999)
Contact: Robert Strausberg, Ph.D.
Email: rcgaps@rmail.nih.gov
Tissue Procurement: ATCC
CDNA Library Preparation: Ling Hong/Rubin Laboratory
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at: image.llnl.gov
Plate: LLCM140 row: h column: 11
High quality sequence start: 4
High quality sequence stop: 782.
Location/Qualifiers
1. 840
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:3348154"
/tissue_type="carcinoma cell line"
/lab_host="NIH_MGC_5"
/note="Organ: cervix; Vector: pOTB7; Site_1: XhoI; Site_2:
EcoRI; CDNA made by oligo-dT priming. Directionally
cloned into EcoRI/XhoI sites using the following 5'
adaptor: GGCACGAG(G). Size-selected >500bp for average
insert size 1.8kb. Library constructed by Ling Hong in
the laboratory of Gerald M. Rubin (University of
California, Berkeley) using ZAP-cDNA synthesis kit
(Stratagene) and Superscript II RT (Life Technologies)."
BASE COUNT 210 a 205 c 225 g 200 t
ORIGIN

```

Query Match		37.4%; Score 579; DB 10; Length 840;
Best Local Similarity		99.1%; Pred. No. 1.7e-155;
Matches		582; Conservative 0; Mismatches 5; Indels 0; Gaps 0;
QY	706	TTTTTCATAGGCTCCGCCCTGACAGACATCACAAAATCGACGCTCAAGTCAGAGGT 765
Db	750	TTTTTCATAGGCTCCGCCCTGACAGACATCACAAAATCGACGCTCAAGTCAGAGGT 691
QY	766	GGGAAACCCGACGAGCTATAAGATACACAGCGCTTTCCCTCCGGAAGTCCCTCGTGC 825
Db	690	GGGAAACCCGACGAGCTATAAGATACACAGCGCTTTCCCTCCGGAAGTCCCTCGTGC 631
QY	826	GCTCTCCTGTTCGACCCCTGCGCTTACCGGATACCTGTCGCGCTTTCTCCCTCCGGAA 885
Db	630	GCTCTCCTGTTCGACCCCTGCGCTTACCGGATACCTGTCGCGCTTTCTCCCTCCGGAA 571
QY	886	GCCTGGCGCTTCTCAATGCTCAGCTGTAGGTATCTCAGTTGGGTAGGTCTTCGCT 945
Db	570	GCCTGGCGCTTCTCATAGCTCAGCTGTAGGTATCTCAGTTGGGTAGGTCTTCGCT 511
QY	946	CCAAGCTGGGCTGTGTGACGAACCCCGCTTCAGCCCGACCGCTGCGCTTATCCGGTA 1005
Db	510	CCAAGCTGGGCTGTGTGACGAACCCCGCTTCAGCCCGACCGCTGCGCTTATCCGGTA 451
QY	1006	ACTATCGCTTGTAGTCCAAACCCGGTAAGACACAGCTTATGCGCACTGGCAGCCACTG 1065
Db	450	ACTATCGCTTGTAGTCCAAACCCGGTAAGACACAGCTTATGCGCACTGGCAGCCACTG 391
QY	1066	GTAACAGGATTACAGCGAGGTATGTAGGCGGTGCTACAGAGTTCTGAAGTGGTGC 1125
Db	390	GTAACAGGATTACAGCGAGGTATGTAGGCGGTGCTACAGAGTTCTGAAGTGGTGC 331
QY	1126	CTAACTACGGCTACACTAGAAGACAGTATTTGGTATCTCGCTCTGCTGAAGCCAGTTA 1185
Db	330	CTAACTACGGCTACACTAGAAGACAGTATTTGGTATCTCGCTCTGCTGAAGCCAGTTA 271
QY	1186	CTTCGGAAAGAGTTGGTAGTCTTTGATCCGGCAAAACACCCGCTGGTAGCGGTG 1245
Db	270	CTTCGGAAAGAGTTGGTAGTCTTTGATCCGGCAAAACACCCGCTGGTAGCGGTG 211
QY	1246	GTTTTTTTGTTTTCAGCAGCAGATTACGGCGCAGAAAAAAGGATCT 1292
Db	210	GTTTTTTTGTTTTCAGCAGCAGATTACGGCGCAGAAAAAAGGATCT 164
RESULT 8		
A0876119/c		
LOCUS		
DEFINITION		A0876119 784 bp DNA GSS 08-NOV-1999
ACCESSION		V133E1 mtN-3xHA/lacZ Insertion Library, strain Y2278 Saccharomyces
VERSION		cerevisiae genomic 5', DNA sequence.
KEYWORDS		A0876119.1 GI:6288363
SOURCE		GSS.
ORGANISM		baker's yeast.
REFERENCE		Saccharomyces cerevisiae
AUTHORS		Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
		Saccharomycetales; Saccharomycetaceae; Saccharomycetes.
		1 (bases 1 to 784)
		Ross-Macdonald, P., Roemer, T., Coelho, P. S. R., Agarwal, S., Kumar, A.,
		desRages, S. A., Cheung, K. H., Sheehan, A., Symoniat, D., Jansen, R.,
		Umsky, L., Heidtman, M., Nelson, K., Iwasaki, H., Kanada, D., Lugo, R.,
		Hager, K., Miller, P., Roeder, G. S., and Snyder, M.
		Large-Scale Analysis of the Yeast Genome by Transposon Tagging and
		Gene Disruption
		Unpublished (1999)
		Contact: Kumar A
		Michael Snyder, Dept. of Mol. Cell. and Dev. Biology
		Yale University
		P.O. Box 208103, New Haven, CT 06520-8103, USA
		Tel: 203 432 9949
		Fax: 203 432 6161
		Email: anuj.kumaryale.edu
		te of mtN-3xHA/lacZ insertion.
TITLE		
JOURNAL		
COMMENT		
Seq primer: GGCCTCTTCTTTTGGAGTAC		
Class: transposon-tagged.		
Location/Qualifiers		
1. .784		
/organism="Saccharomyces cerevisiae"		
/strain="Y2278 - S288C background, cir(0) rho(0)"		
/db_xref="taxon:4932"		
/clone_lib="mtN-3xHA/lacZ Insertion Library, strain Y2278"		
/lab_host="E. coli"		
/note="Vector: pHS6-Sal; A yeast genomic DNA library		
without 2 micron or mitochondrial DNA was prepared in		
pHS6-Sal; genomic DNA was size-fractionated (DNA of		
roughly 2-3 kb in length) prior to cloning. This library		
was subsequently mutagenized with a mtN-3xHA/lacZ		
minitransposon containing lacZ, URA3, and tet resistance.		
BASE COUNT		187 a 200 c 210 g 186 t 1 others
ORIGIN		
Query Match		37.2%; Score 576.2; DB 13; Length 784;
Best Local Similarity		97.7%; Pred. No. 1.1e-154;
Matches		584; Conservative 0; Mismatches 14; Indels 0; Gaps 0;
QY	695	GTTCGCTGGGCTTTTCCATAGGCTCCGCCCTGACGAGCATCAGAAAATCGAGCTC 754
Db	784	GTTCGCTGGGCTTTTCCATAGGCTCCGCCCTGACGAGCATCAGAAAATCGAGCTC 725
QY	755	AAGTCAGAGTGGCGAACCAGACAGACTATAAGATACACAGCGCTTCCCTCGGAAG 814
Db	724	AAGTCAGAGTGGCGAACCAGACAGACTATAAGATACACAGCGCTTCCCTCGGAAG 665
QY	815	CTCCCTCGTCCGCTCTCCGCTTCCGCTTACCGGATACCTGTCGCGCTTCT 874
Db	664	CTCCCTCGTCCGCTCTCCGCTTCCGCTTACCGGATACCTGTCGCGCTTCT 605
QY	875	CCCTTCGGGAAGCGTGGCGCTTTCTCAAGCTCAGCGTGTAGGTATCTAGTTCGGTGA 934
Db	604	CCCTTCGGGAAGCGTGGCGCTTTATCATAGCTCAGCGTGTAGGTATCTAGTTCGGTGA 545
QY	935	GGTCGCTCCGCTCCAAAGCTGGGCTGTGTCAGCAACCCCGCTTCAGCCGACCTCGC 994
Db	544	GGTCGCTCCGCTCCAAAGCTGGGCTGTGTCAGCAACCCCGCTTCAGCCGACCTCGC 485
QY	995	CTTATCCGCTTAAGTCTTGGAGTCCAAACCCGGTAAGACAGACTTATCGCCACTGGC 1054
Db	484	CTTATCCGCTTAAGTCTTGGAGTCCAAACCCGGTAAGACAGACTTATCGCCACTGGC 425
QY	1055	AGCAGCCACTGGTAACAGGATTAGCAGAGCGAGGTATGTAGGCGGTGCTACAGATTCT 1114
Db	424	AGCAGCCACTGGTAACAGGATTAGCAGAGCGAGGTATGTAGGCGGTGCTACAGATTCT 365
QY	1115	GAAGTGGTGGCTTAAGTCTTACGCTTACGCTACACTAGAAGACAGTATTTGGTATCTCGCTCTGCT 1174
Db	364	GAAGTGGTGGCTTAAGTCTTACGCTTACGCTTACACTAGAAGACAGTATTTGGTATCTCGCTCTGCT 305
QY	1175	GAAGCCAGTTTACCTTCGGAAGAAAGAGTTGGTACTCTTGTATCCGGCAACAAACACCGC 1234
Db	304	GAAGCCAGTTTACCTTCGGAAGAAAGAGTTGGTACTCTTGTATCCGGCAACAAACACCGC 245
QY	1235	TGTTAGCGGTGGTTTTTTTTCGTAAGCAGCAGATTACCGCGCAGAAAAAAGGATCT 1292
Db	244	TGTTAGCGGTGGTTTTTTTTCGTAAGCAGCAGATTACCGCGCAGAAAAAAGGATCT 187
RESULT 9		
A0876011/c		
LOCUS		
DEFINITION		A0876011 795 bp DNA GSS 08-NOV-1999
ACCESSION		V132B5 mtN-3xHA/lacZ Insertion Library, strain Y2278 Saccharomycetes
VERSION		cerevisiae genomic 5', DNA sequence.
KEYWORDS		A0876011
SOURCE		A0876011.1 GI:6288255
ORGANISM		GSS.

SOURCE
ORGANISM
baker's yeast.
Saccharomyces cerevisiae
Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
Saccharomycetales; Saccharomycetaceae; Saccharomycetes.
REFERENCE
AUTHORS
1 (bases 1 to 795)
Ross-Macdonald, P., Roemer, T., Coelho, P.S.R., Agarwal, S., Kumar, A.,
desEtages, S.A., Cheung, K.-H., Sheehan, A., Symoniatidis, D., Jansen, R.,
Umansky, L., Heldtman, M., Nelson, K., Iwasaki, H., Kanada, D., Lugo, R.,
Hager, K., Miller, P., Roeder, G.S., and Snyder, M.
TITLE
Large-Scale Analysis of the Yeast Genome by Transposon Tagging and
Gene Disruption
JOURNAL
COMMENT
Unpublished (1999)
Contact: Kumar A
Michael Snyder, Dept. of Mol. Cell. and Dev. Biology
Yale University
P.O. Box 208103, New Haven, CT 06520-8103, USA
Tel: 203 432 9949
Fax: 203 432 6161
Email: anuj.kumar@yale.edu
te of mTn-3xHA/lacZ insertion.
Seq primer: GGCTTCTCTTTTGGAGTAC
Class: transposon-tagged.
FEATURES
Source
Location/Qualifiers
1..795
/organism="Saccharomyces cerevisiae"
/strain="Y2278 - S288C background, cir(0) rho(0)"
/db_xref="taxon:4932"
/clone_lib="mTn-3xHA/lacZ Insertion Library, strain Y2278"
/lab_host="E. coli"
/note="Vector: pHS56-Sal; A yeast genomic DNA library
without 2 micron or mitochondrial DNA was prepared in
pHS56-Sal; genomic DNA was size-fractionated (DNA of
roughly 2-3 kb in length) prior to cloning. This library
was subsequently mutagenized with a mTn-3xHA/lacZ
minitransposon containing lacZ, URA3, and tet resistance."

BASE COUNT 191 a 203 c 212 g 188 t 1 others

Query Match 37.1%; Score 573.6; DB 13; Length 795;
Best Local Similarity 99.1%; Pred. No. 6e-154; 5; Indels 0; Gaps 0;
Matches 576; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 712 ATAGGCTCCGCCGCCCTGACGAGCATCACAAAATCGACGCTCAAGTCAGAGGTGGCGAA 771
DB 795 ATAGGCTCCGCCGCCCTGACGAGCATCACAAAATCGACGCTCAAGTCAGAGGTGGCGAA 736
QY 772 ACCGACAGACTATAAAGATACACAGCGTTTCCCTCGGAAGTCCCTCGTGGCGTCTC 831
DB 735 ACCGACAGACTATAAAGATACACAGCGTTTCCCTCGGAAGTCCCTCGTGGCGTCTC 676
QY 832 CTGTTCCGACCCCTGCGCTTACCGGATACCTGTCGCGCTTTCCTCCCTCGGAGCGTGG 891
DB 675 CTGTTCCGACCCCTGCGCTTACCGGATACCTGTCGCGCTTTCCTCCCTCGGAGCGTGG 616
QY 892 CGCTTTCTCAATGCTACCGCTGTAGGTATCTCAGTTCGGTGTAGGTGCTCGCTCCAAGC 951
DB 615 CGCTTTCTCAATGCTACCGCTGTAGGTATCTCAGTTCGGTGTAGGTGCTCGCTCCAAGC 556
QY 952 TGGGCTGTGTGACGAACCCCGCTTACGCGCCGACCGCTCGCTTATCCGGTAACTATC 1011
DB 555 TGGGCTGTGTGACGAACCCCGCTTACGCGCCGACCGCTCGCTTATCCGGTAACTATC 496
QY 1012 GTCTTGAGTCCACCCGGTAAACACACACTTATCGCCACATGCGGACGACACTGGTAACA 1071
DB 495 GTCTTGAGTCCACCCGGTAAACACACACTTATCGCCACATGCGGACGACACTGGTAACA 436
QY 1072 GGATTACGACGCGAGTATGTAGCGGCTGCTACAGAGTTCTTCAAGTGTGGCTAACT 1131
DB 435 GGATTACGACGCGAGTATGTAGCGGCTGCTACAGAGTTCTTCAAGTGTGGCTAACT 376
QY 1132 ACGGCTACACTAGAAAGACAGTATTTGGTATCTGCGCTCTGCTGAAGCCAGTTACCTTCG 1191

Db 375 ACGGCTACACTAGAAAGACAGTATTTGGTATCTGCGCTCTGCTGAAGCCAGTTACCTTCG 316
QY 1192 GAAAAAGAGTTGGTAGCTCTGTATCCGGCAACAAACACCGCTGTAGCGGTGTTTTT 1251
Db 315 GAAAAAGAGTTGGTAGCTCTGTATCCGGCAACAAACACCGCTGTAGCGGTGTTTTT 256
QY 1252 TTGTTTTCACAGCAGCAGATTCACGCGCAGAAAAAAGGATCT 1292
Db 255 TTGTTTTCACAGCAGCAGATTCACGCGCAGAAAAAAGGATCT 215

RESULT 10
AJ281661
LOCUS
DEFINITION
4A3A-P8G10-F Anopheles gambiae immune competent 4A3A Anopheles
gambiae cDNA clone 4A3A-P8G10, mRNA sequence.
AJ281661
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
Anopheles gambiae
African malaria mosquito.
Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
Pterygota; Neoptera; Endopterygota; Diptera; Nematocera; Culicoidea
; Anopheles.
REFERENCE
AUTHORS
1 (bases 1 to 617)
Dimopoulos G., Casavant, T.L., Chang, S., Scheetz, T., Roberts, C.,
Donohue, M., Schultz, J., Benes, V., Bork, P., Ansorge, W., Soares, M.B.
and Kafatos, F.C.
TITLE
Anopheles gambiae pilot gene discovery project: identification of
mosquito innate immunity genes from expressed sequence tags
generated from immune-competent cell lines
Proc. Natl. Acad. Sci. U.S.A. 97 (12), 6619-6624 (2000).
JOURNAL
MEDLINE
COMMENT
20300950
Contact: Dimopoulos G
Fotis C. Kafatos laboratory
European Molecular Biology Laboratory
Meyerhofstrasse 1, 69117 Heidelberg, Germany.
FEATURES
Source
Location/Qualifiers
1..617
/organism="Anopheles gambiae"
/strain="4A r/r"
/db_xref="taxon:7165"
/clone_lib="4A3A-P8G10"
/cell_line="immune competent 4A3A"
/lab_host="E. coli DH10B"
/note="Vector: pT73D-Pac (Pharmacia) with a modified
polylinker; Site_1: EcoRI; Site_2: NotI; sequenced from
forward priming site which reads from the 3' end of the
cDNA. The 4A3A is a directionally cloned and normalized
cDNA library that was constructed from the 4A3A cell line
oligo-T primed cDNA according to: Bonaldo, Lennon & Soares
(1996) : Normalization and Subtraction: Two approaches to
Facilitate Gene Discovery, Genome Research 6, 791-806."

BASE COUNT 136 a 175 c 162 g 144 t

Query Match 36.5%; Score 565.2; DB 10; Length 617;
Best Local Similarity 99.3%; Pred. No. 1.5e-151;
Matches 578; Conservative 0; Mismatches 3; Indels 1; Gaps 1;

QY 690 GCCCGTGTGGCGTTTTCATAGGCTCCGCCCTCGACGAGCATCACAAAATCGA 749
Db 37 GCCCGTGTGGCGTTTTCATAGGCTCCGCCCTCGACGAGCATCACAAAATCGA 96
QY 750 CGCTCAAGTCAGAGTGGCGAAACCCGACAGGACTATAAAGATACACAGCGCTTCCCGCT 809
Db 97 CGCTCAAGTCAGAGTGGCGAAACCCGACAGGACTATAAAGATACACAGCGCTTCCCGCT 156
QY 810 GGAAGCTCCCTCGTGGCGTCTCTCTGTTCCGACCTGCGCTTACCGGATACCTGCGCC 869

Db 157 GGAAGCTCCCTCGTGGCTCTCTCTGTTCCGACCCCTTACCGGATACCTGTCGCGC 216

QY 870 TTTCTCCCTTCGGGAAGCTGGCGCTTCTCAATGCTCAGCTGTAGGTATCTCAGTTGG 929

Db 217 TTTCTCCCTTCGGGAAGCTGGCGCTTCTCAATGCTCAGCTGTAGGTATCTCAGTTGG 276

QY 930 GTCTAGTGTCTGCTGCCAAGCTGGGCTGTGTGACGAACCCCGCTTTCAGCCGACCGC 989

Db 277 GTGTAGTGTCTGCTGCCAAGCTGGGCTGTGTGACGAACCCCGCTTTCAGCCGACCGC 336

QY 990 TGGCCCTTATCCGGTAACATATCGTCTTGAGTCCAAACCCCGTAAGACACGACTTATCGCCA 1049

Db 337 TGGCCCTTATCCGGTAACATATCGTCTTGAGTCCAAACCCCGTAAGACACGACTTATCGCCA 396

QY 1050 CTGGCAGCAGCCACTGGTAACAGGATTAGCAGAGCAGGTATGTAGGCGGTGCTACAGAG 1109

Db 397 CTGGCAGCAGCCACTGGTAACAGGATTAGCAGAGCAGGTATGTAGGCGGTGCTACAGAG 456

QY 1110 TTTCTCAAGTGGTGGCTTAACAGGCTTACAGGCTACACTAGAGGACAGTATTTGATCTCGCT 1169

Db 457 TTTCTCAAGTGGTGGCTTAACAGGCTTACAGGCTACACTAGAGGACAGTATTTGATCTCGCT 516

QY 1170 CTGCTGAAGCCAGTTACCTTCGGAAGAGTGGTAGCTTTGATCCGCGCAACAAACC 1229

Db 517 CTGCTGAAGCCAGTTACCTTCGGAAGAGTGGTAGCTTTGATCCGCGCAACAAACC 576

QY 1230 ACCGCTGGTAGCGGTGGTGTGTTTGTGTCGACGACGAGATT 1271

Db 577 ACCGCTGGTAGC-GTGGTGTGTTGTTGTCGACGACGAGATT 617

RESULT 11

AL044178 571 bp mRNA EST 29-FEB-2000

LOCUS DKF2p434P0828_s1.434 (synonym: htes3) Homo sapiens cDNA clone

DEFINITION DKF2p434P0828 3', mRNA sequence.

ACCESSION AL044178

VERSION AL044178.3 GI:5935954

KEYWORDS EST.

SOURCE human.

ORGANISM Homo sapiens

REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

AUTHORS Bloecker, H., Boecker, M., Brandt, P., Mewes, H.W., Gassenhuber, J. and Wiemann, S.

TITLE EST (Bloecker, et al.)

JOURNAL Unpublished (1999)

COMMENT On Jul 9, 1999 this sequence version replaced gi:5866789.

Contact: Bloecker H

MPIS

Am Klopferpitz 18a D-82152 Martinsried, Germany

This is the 3' sequence of the clone insert

Clone from S. Wiemann, Molecular Genome Analysis, German Cancer Research Center (DKFZ); Email s.wiemann@dkfz-heidelberg.de;

sequenced by GBF (National Research Centre for Biotechnology Ltd., Braunschweig/Germany) within the cDNA sequencing consortium of the German Genome Project.

No r1 sequence available.

This clone (DKF2p434P0828) is available at the RZPD in Berlin.

Please contact the RZPD: Ressourcenzentrum, Heubnerweg 6, 14059 Berlin-Charlottenburg, GERMANY; Email: clone@rzpd.de.

Location/Qualifiers

1. 571

/organism="Homo sapiens"

/db_xref="taxon:9606"

/clone="DKF2p434P0828"

/clone_lib="434 (synonym: htes3)"

/tissue_type="testis"

/dev_stage="adult"

/lab_host="DH108"

/note="Vector: pSport1; Site_1: NotI; Site_2: SalI"

122 a 168 c 149 g 132 t

BASE COUNT

ORIGIN

Query Match 35.9%; Score 555.2; DB 10; Length 571;

Best Local Similarity 99.5%; Pred. No. 1.1e-148;

Matches 557; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 690 GCGCGGTGCTGGCGCTTTTCCATAGGCTCCGCCCGCTTACGAGCATCACAATAATCGA 749

Db 12 GCGCGGTGCTGGCGCTTTTCCATAGGCTCCGCCCGCTTACGAGCATCACAATAATCGA 71

QY 750 CGCTCAAGTCAAGGTGGGAAACCCGACAGGACTATAAGATACCAAGCGTTTCCCGCT 809

Db 72 CGCTCAAGTCAAGGTGGGAAACCCGACAGGACTATAAGATACCAAGCGTTTCCCGCT 131

QY 810 GGAAGCTCCCTCGTGGCTCTCTGTTCCGACCCCTTACCGGATACCTGTCCGCG 869

Db 132 GGAAGCTCCCTCGTGGCTCTCTGTTCCGACCCCTTACCGGATACCTGTCCGCG 191

QY 870 TTTCTCCCTTCGCGAAGCGTGGCGCTTTCTCAATGCTACGCTGTAGGTATCTCAGTTGG 929

Db 192 TTTCTCCCTTCGCGAAGCGTGGCGCTTTCTCAATGCTACGCTGTAGGTATCTCAGTTGG 251

QY 930 GTGTAGGTGCTGGCTCCAAAGCTGGCGTGTGTGACGAACCCCGCTTACGCGCGACCGC 989

Db 252 GTGTAGGTGCTGGCTCCAAAGCTGGCGTGTGTGACGAACCCCGCTTACGCGCGACCGC 311

QY 990 TCGCGCTTATCCGGTAACATATCGTCTTGAGTCCAAACCCCGTAAGACACGACTTATCGCCA 1049

Db 312 TCGCGCTTATCCGGTAACATATCGTCTTGAGTCCAAACCCCGTAAGACACGACTTATCGCCA 371

QY 1050 CTGGCAGCAGCCACTGGTAACAGGATTAGCAGAGCAGGTATGTAGGCGGTGCTACAGAG 1109

Db 372 CTGGCAGCAGCCACTGGTAACAGGATTAGCAGAGCAGGTATGTAGGCGGTGCTACAGAG 431

QY 1110 TTTCTCAAGTGGTGGCTTAACGCTACGCTACAGGACGACTATTTGTTATCTCGCT 1169

Db 432 TTTCTCAAGTGGTGGCTTAACGCTACGCTACAGGACGACTATTTGTTATCTCGCT 491

QY 1170 CTGCTGAAGCCAGTTACCTTCGGAAGAGTGGTAGCTTTGATCCGCGCAACAAACC 1229

Db 492 CTGCTGAAGCCAGTTACCTTCGGAAGAGTGGTAGCTTTGATCCGCGCAACAAACC 551

QY 1230 ACCGCTGGTAGCGGTGGTGT 1249

Db 552 ACCGCTGGTAGCGGTGT 571

RESULT 12

AJ281320 579 bp mRNA EST 30-JUN-2000

LOCUS 4A3A-PIH1-F Anopheles gambiae immune competent 4A3A Anopheles

DEFINITION gambiae cDNA clone 4A3A-PIH1, mRNA sequence.

ACCESSION AJ281320

VERSION AJ281320.1 GI:6929201

KEYWORDS EST.

SOURCE African malaria mosquito.

ORGANISM Anopheles gambiae

Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta; Pterygota; Neoptera; Endopterygota; Diptera; Nematocera; Culicoidae

REFERENCE 1 (bases 1 to 579)

AUTHORS Dimopoulos, G., Casavant, T.L., Chang, S., Scheetz, T., Roberts, C., Donohue, M., Schultz, J., Benes, V., Bork, P., Ansoorge, W., Soares, M.B. and Kafatos, F.C.

TITLE Anopheles gambiae pilot gene discovery project: identification of mosquito innate immunity genes from expressed sequence tags generated from immune-competent cell lines

JOURNAL Proc. Natl. Acad. Sci. U.S.A. 97 (12), 6619-6624 (2000)

MEDLINE 20300950

COMMENT Contact: Dimopoulos G

Fortis C. Kafatos laboratory

European Molecular Biology Laboratory


```

FEATURES
  source
    Meyerhofstrasse 1, 69117 Heidelberg, Germany.
    1..579
    /organism="Anopheles gambiae"
    /strain="4A r/r"
    /db_xref="taxon:7165"
    /clone="4A3A-PIH1"
    /cell_line="immune competent 4A3A"
    /lab_host="E. coli DH10B"
    /note="Vector: pT773D-Pac (Pharmacia) with a modified
polylinker; Site_1: EcoRI; Site_2: NotI; sequenced from
forward priming site which reads from the 3' end of the
cDNA. The 4A3A is a directionally cloned and normalized
cDNA library that was constructed from the 4A3A cell line
oligo-T primed cDNA according to: Bonaldo, Lennon & Soares
(1996) : Normalization and Subtraction: Two approaches to
Facilitate Gene Discovery, Genome Research 6, 791-806."
BASE COUNT 120 a 167 c 151 g 137 t 4 others
ORIGIN
Query Match 35.8%; Score 554; DB 10; Length 579;
Best Local Similarity 98.1%; Pred. No. 2.4e-148;
Matches 557; Conservative 3; Mismatches 8; Indels 0; Gaps 0;
QY 690 GCGCGTGTGCTGGCTTTTCCATAGCTCCGCCCTGACGAGCATCAAAAAATCGA 749
Db 5 GCGCGTGTGCTGGCTTTTCCATAGCTCCGCCCTGACGAGCATCAAAAAATCGA 64
QY 750 CGCTCAAGTCAGAGTGGCGAAACCCGACAGGACTATAAAGATACACGAGGTTTCCCGCT 809
Db 65 CGCTCAAGTCAGAGTGGCGAAACCCGACAGGACTATAAAGATACACGAGGTTTCCCGCT 124
QY 810 GGAAGTCCCTCGTGGCTCTCTGTTTCGACCGCTGCGGTTACCGGATACCTGTCCGCC 869
Db 125 GGAAGTCCCTCGTGGCTCTCTGTTTCGACCGCTGCGGTTACCGGATACCTGTCCGCC 184
QY 870 TTCTCTCCCTCGGAGCGTGGCGCTTCTCAATGCTCAGCTGATGATCTCAGTTCG 929
Db 185 TTCTCTCCCTCGGAGCGTGGCGCTTCTCAATGCTCAGCTGATGATCTCAGTTCG 244
QY 930 GTGTAGTCTGCTCGCTCCAAAGCTGGGCTGTGTGCACGAACCCCGCTTACGCCGACCGC 989
Db 245 GTGTAGTCTGCTCGCTCCAAAGCTGGGCTGTGTGCACGAACCCCGCTTACGCCGACCGC 304
QY 990 TGCCTCTATCCGCTAACTATCGCTTTAGTCCAAACCCGGTAAGACACGACTTATCGCCA 1049
Db 305 TGCCTCTATCCGCTAACTATCGCTTTAGTCCAAACCCGGTAAGACACGACTTATCGCCA 364
QY 1050 CTGCACGACGCTGTATACAGATTAGCAGACGAGGATGTAGCGGCTGTACAGAG 1109
Db 365 CTGCACGACGCTGTATACAGATTAGCAGACGAGGATGTAGCGGCTGTACAGAG 424
QY 1110 TTCTTGAAGTGGTGGCTTAACAGGCTACACTAGAGGACAGTATTTGATCTCGCT 1169
Db 425 TTCTTGAAGTGGTGGCTTAACAGGCTACACTAGAGGACAGTATTTGATCTCGCT 484
QY 1170 CTGCTGAAGCCAGTACCTCTCGGAAAAGAGTTGGTAGCTCTTGATCCGCAACAAC 1229
Db 485 CTGCTGAAGCCAGTACCTCTCGGAAAAGAGTTGGTAGCTCTTGATCCGCAACAAC 544
QY 1230 ACCGCTGGTAGCGGTGTTTTTTTGT 1257
Db 545 ACCGCTGGTAGCGGTGTTTTTTTGT 572

RESULT 13
AQ875988/c 794 bp DNA GSS 08-NOV-1999
LOCUS
DEFINITION V131h3 mtN-3xHA/lac2 Insertion Library, strain Y2278 Saccharomyces
cerevisiae genomic 5', DNA sequence.
ACCESSION AQ875988

```

```

VERSION AQ875988.1 GI:6288232
KEYWORDS GSS.
SOURCE baker's yeast.
ORGANISM Saccharomyces cerevisiae
REFERENCE Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
AUTHORS Saccharomycetales; Saccharomycetaceae; Saccharomycetes.
1 (bases 1 to 794)
Ross-Macdonald, P., Roemer, T., Coelho, P. S. R., Agarwal, S., Kumar, A.,
desRages, S. A., Cheung, K. H., Sheehan, A., Symoniatidis, D., Jansen, R.,
Mansky, L., Heidtman, M., Nelson, K., Iwasaki, H., Kanada, D., Lugo, R.,
Hager, K., Miller, P., Roeder, G. S., and Snyder, M.
Large-Scale Analysis of the Yeast Genome by Transposon Tagging and
Gene Disruption
Unpublished (1999)
Contact: Kumar A
Michael Snyder, Dept. of Mol. Cell. and Dev. Biology
Yale University
P.O. Box 208103, New Haven, CT 06520-8103, USA
Tel: 203 432 9949
Fax: 203 432 6161
Email: anuj.kumare@yale.edu
te of mtN-3xHA/lac2 insertion.
Seq primer: GGCCTCTCTCTTTTGAAGTAC
Class: transposon-tagged.
FEATURES
  source
    1..794
    /organism="Saccharomyces cerevisiae"
    /strain="Y2278 - S288C background, cir(0) rho(0)"
    /db_xref="taxon:4932"
    /clone_lib="mtN-3xHA/lac2 Insertion Library, strain Y2278"
    /lab_host="E. coli"
    /note="Vector: pHS56-Sal; A yeast genomic DNA library
without 2 micron or mitochondrial DNA was prepared in
pHS56-Sal; genomic DNA was size-fractionated (DNA of
roughly 2-3 kb in length) prior to cloning. This library
was subsequently mutagenized with a mtN-3xHA/lac2
minitransposon containing lacZ, URA3, and tet resistance."
BASE COUNT 186 a 193 c 220 g 191 t 4 others
ORIGIN
Query Match 35.8%; Score 554; DB 13; Length 794;
Best Local Similarity 98.3%; Pred. No. 2.6e-148;
Matches 579; Conservative 0; Mismatches 8; Indels 2; Gaps 2;
QY 706 TTTTCCATAGGCTCCGCCCTGACGAGCATCAAAAAATCGAGTCAAGTCAGAGT 765
Db 788 TTTTCCATAGGCTCCGCCCTGACGAGCATCAAAAAATCGAGTCAAGTCAGAGT 729
QY 766 GCGAAACCGGACA-GGACTATAAAGATA-CCAGGCGTTTCCCGCTGGAAGCTCCCTCGT 823
Db 728 GCGAAACCGGACAGGCGCTATAAAGATATCCAGGCGTTTCCCGCTGGAAGCTCCCTCGT 669
QY 824 GCGCTCTCTGTTCCGACCTGCGCTTACCGGATACCTGTCGCCCTTCTCCCTTCGGG 883
Db 568 GCGCTCTCTGTTCCGACCTGCGCTTCCGGAATACCTGTCGCCCTTCTCCCTTCNGG 609
QY 884 AAGCGTGGCGCTTTCTCAATGCTCACGCTGTAGGTATCTCAGTTCGGTGTAGTGTG 943
Db 608 AAGCGTGGCGCTTTCTCATAGCTCACGCTGTAGGTATCTCAGTTCGGTGTAGTGTG 549
QY 944 CTCGAAGCTGGCTGTGTGCACGAACCCCGCTTACGCCGACCGCTGCGGCTTATCCGG 1003
Db 548 CTCGAAGCTGGCTGTGTGCACGAACCCCGCTTACGCCGACCGCTGCGGCTTATCCGG 489
QY 1004 TAACATCTGCTTTCAGTCCAAACCGGTAAGACAGACTTATCGCCACTGGCAGCAGCAC 1063
Db 488 TAACATCTGCTTTCAGTCCAAACCGGTAAGACAGACTTATCGCCACTGGCAGCAGCAC 429
QY 1064 TGGTAACAGGATTAGCAGACGAGGATGTAGGCGGCTGTACAGAGTCTTGAAGTGGTG 1123
Db 428 TGGTAACAGGATTAGCAGACGAGGATGTAGGCGGCTGTACAGAGTCTTGAAGTGGTG 369

```

```

QY 1124 GCCTAACTACGGCTACACTAGAGACAGTATTGTGATCTCGCTCTGCTGAAGCCACT 1183
|||||
Db 368 GCTACTACGGCTACACTAGAGACAGTATTGTGATCTCGCTCTGCTGAAGCCACT 309
|||||
QY 1184 TACCTTCGGAAGAGTGGTAGCTCTTGTATCGCGCAACAAACACCGCTGGTAGCGG 1243
|||||
Db 308 TACCTTCGGAAGAGTGGTAGCTCTTGTATCGCGCAACAAACACCGCTGGTAGCGG 249
|||||
QY 1244 TGGTTTTTTTGTTCGAACGACAGATTACGCGCGCAAAAAAGGATCT 1292
|||||
Db 248 TGGTTTTTTTGTTCGAACGACAGATTACGCGCGCAAAAAAGGATCT 200
|||||

RESULT 14
AV735664 AV735664 615 bp mRNA EST 17-OCT-2000
LOCUS
DEFINITION AV735664 CB Homo sapiens cDNA clone CBNAME07 5', mRNA sequence.
ACCESSION AV735664
VERSION AV735664.1 GI:10853245
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 615)
AUTHORS Zhang, Q., Ye, M., Wu, X., Gu, J., Huang, Q., Zhou, J., Shen, Y., Han, Z.,
Chen, S., Mao, M. and Chen, Z.
TITLE Homo sapiens CB library cDNA clones
JOURNAL Unpublished (2000)
COMMENT Contact: Zhu Chen
Shanghai Institute of Hematology, Rui-Jin Hospital
197 Rui-Jin II Road, Shanghai 200025, P. R. China
Tel: 86-21-64740490
Fax: 86-21-64743206
Email: mbshiems.scn.sh.cn
This clone is available at Shanghai Hematology Institute in
Shanghai.
Chinese National Human Genome Center at Shanghai
351 Guo Shoujing Road, Zhangjiang Hi- Tech Park, Pudong.
Location/Qualifiers
FEATURES
source
1..615
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="CBNAME07"
/clone_lib="CB"
/tissue_type="cord blood"
/cell_type="CD34+ hematopoietic stem/progenitor cell"
/lab_host="BM25.8"
/note="Vector: pBluescript; Site_1: EcoRI; The insert is
cloned randomly with the EcoRI digestion"
BASE COUNT 129 a 176 c 159 g 147 t
ORIGIN
Query Match 35.8%; Score 553.2; DB 10; Length 615;
Best Local Similarity 97.5%; Pred. No. 4.1e-148;
Matches 591; Conservative 0; Mismatches 12; Indels 3; Gaps 3;

QY 688 CGCGCGCTGTGGGTTTCCATAGGCTCGCGCCCTGACGAGCATCAAAAATC 747
|||||
Db 7 CGCGCGCTGTGGGTTTCCATAGGCTCGCGCCCTGACGAGCATCAAAAATC 66
|||||
QY 748 GAGCGTCAAGTCAGAGTGGGAAACCCGACAGGACTATAAGATACCGGCTTTCCCC 807
|||||
Db 67 GAGCGTCAAGTCAGAGTGGGAAACCCGACAGGACTATAAGATACCGGCTTTCCCC 126
|||||
QY 808 CTGGAAGCTCCCTCGTGCCTCTCTCTGTTCCGACCTGCGGTTACCGATACCTGTCG 867
|||||
Db 127 CTGGAAGCTCCCTCGTGCCTCTCTCTGTTCCGACCTGCGGTTACCGATACCTGTCG 186
|||||
QY 868 CTTTCTCCCTTCGGGAACGCTGGCGCTTTCATGCTCAGCTGTAGTATCTCAGTT 927
|||||

```

```

Db 187 CTTTCTCCCTTCGGAAGCGTGGCGCTTCTCTATAGCTCAGCTGTAGGTATCTCAGTT 246
QY 928 CGGTGTAGGTCTGCTCCAAAGCTGGGCTGTGTGCACGAACCCCGCTTACGCCGACC 987
|||||
Db 247 CGGTGTAGGTCTGCTCCAAAGCTGGGCTGTGTGCACGAACCCCGCTTACGCCGACC 306
|||||
QY 988 GCTGGCGCTTATCCGGTAACTATCTCTTTAGTCTCAACCCGGTAAAGACACGACTTATCCG 1047
|||||
Db 307 GCTGGCGCTTATCCGGTAACTATCTCTTTAGTCTCAACCCGGTAAAGACACGACTTATCCG 366
|||||
QY 1048 CACTGGCAGCAGCACTGTGTACAGGATTAGCAGCGAGGATGTAGCGGTGTCTACAG 1107
|||||
Db 367 CACTGGCAGCAGCACTGTGTACAGGATTAGCAGCGAGGATGTAGCGGTGTCTACAG 426
|||||
QY 1108 AGTTCTTGAAGTGGGCTAACTACGGCTACACTAGAAGGACAGTATTGGTATCTCGG 1167
|||||
Db 427 AGTTCTTGAAGTGGGCTAACTACGGCTACACTAGAAGGACAGTATTGGTATCTCGG 486
|||||
QY 1168 CTCTGCTGAAGCCAGTTACC-TTCGGAAAAAGAGTTGGTAGCTCTTTGATCCGGCAACAA 1226
|||||
Db 487 CTCTGCTGAAGCCAGTTACC-TTCGGAAAAAGAGTTGGTAGCTCTTTGATCCGGCAACAA 546
|||||
QY 1227 ACCACGGCTGTAGCGGTGTTTGTTCGACGACGACAGATT-ACGCGCAGAAAAA 1285
|||||
Db 547 ACCACGGCTGTAGCGGTGTTTGTTCGACGACGACAGATTAAACGCGNGGATTCA 605
|||||
QY 1286 AGGATC 1291
|||||
Db 606 AGGATC 611
|||||

RESULT 15
AL044364 AL044364 954 bp mRNA EST 29-FEB-2000
LOCUS
DEFINITION DKFZp434C172.sl.434 (synonym: htes3) Homo sapiens cDNA clone
ACCESSION DKFZp434C172.3', mRNA sequence.
VERSION AL044364
KEYWORDS EST.
SOURCE AL044364.1 GI:5432586
human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 954)
AUTHORS Ansorge, W., Benes, V., Krieger, S., Mewes, H.W., Gassenhuber, J. and
Wiemann, S.
TITLE EST (Ansorge, Benes, et al.)
JOURNAL Unpublished (1999)
COMMENT Contact: Ansorge W
MIPS
Am Klopferspitz 18a D-82152 Martinsried, Germany
This is the 3' sequence of the clone insert
Clone from S. Wiemann, Molecular Genome Analysis, German Cancer
Research Center (DKFZ); Email: s.wiemann@dkfz-heidelberg.de;
sequenced by EMBL (European Molecular Biology Laboratories, of the
Heidelberg/Germany) within the cDNA sequencing consortium of the
German Genome Project.
No r1 sequence available.
This clone (DKFZp434C172) is available at the RZPD in Berlin.
Please contact the RZPD: Ressourcenzentrum, Heubnerweg 6, 14059
Berlin-Charlottenburg, GERMANY; Email: clone@rzpd.de.
FEATURES
Location/Qualifiers
1..954
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="DKFZp434C172"
/clone_lib="434 (synonym: htes3)"
/tissue_type="testis"
/dev_stage="adult"
/lab_host="DH10B"
/note="Vector: pSport1; Site_1: NotI; Site_2: SalI"
BASE COUNT 214 a 273 c 253 g 214 t
ORIGIN

```

Query Match		35.6%	Score 550.2;	DB 10;	Length 954;
Best Local Similarity		99.5%	Pred. No. 3.5e-147;		
Matches 552;		Conservative	0;	Mismatches 3;	Indels 0; Gaps 0;
QY	690	GCCGCGTTGCTGGCGTTTTCATAGGCTCCGCCCCCTGACGAGCATCAAAAAATCGA	749		
Db	400	GCCGCGTTGCTGGCGTTTTCATAGGCTCCGCCCCCTGACGAGCATCAAAAAATCGA	459		
QY	750	CGCTCAAGTCAGAGTGGCGAACCACAGGACTATAAGATACCAAGCGTTTCCCCCT	809		
Db	460	CGCTCAAGTCAGAGTGGCGAACCACAGGACTATAAGATACCAAGCGTTTCCCCCT	519		
QY	810	GGAAGTCCCTCGTGGCGCTCCTGTTCGACCTGCGCTTACCGGATACCTGTCCGCC	869		
Db	520	GGAAGTCCCTCGTGGCGCTCCTGTTCGACCTGCGCTTACCGGATACCTGTCCGCC	579		
QY	870	TTTCTCCCTTCGGGAGCGTGGCGCTTCTCAATGCTACGCTGTAGGTATCTCAGTTGC	929		
Db	580	TTTCTCCCTTCGGGAGCGTGGCGCTTCTCATAGCTACGCTGTAGGTATCTCAGTTGC	639		
QY	930	GTGTAGGTCGTTCCGCTCCAAGCTGGGCTGTGTGCAGCAACCCCGTTACGCCGACCGC	989		
Db	640	GTGTAGGTCGTTCCGCTCCAAGCTGGGCTGTGTGCAGCAACCCCGTTACGCCGACCGC	699		
QY	990	TGGCGCTTATCCGGTAACCTATCGTCTTGAGTCCAAACCCGGTAAGACACGACTTATCGCCA	1049		
Db	700	TGGCGCTTATCCGGTAACCTATCGTCTTGAGTCCAAACCCGGTAAGACACGACTTATCGCCA	759		
QY	1050	CTGGCAGCAGCCACTGGTAACAGGATTAGCAGAGCGAGGTATGTAGGCGGTGCTACAGAG	1109		
Db	760	CTGGCAGCAGCCACTGGTAACAGGATTAGCAGAGCGAGGTATGTAGGCGGTGCTACAGAG	819		
QY	1110	TTCTTGAAGTGGTGGCTTAACCTACGGCTACACTAGAGGACAGTATTTGGTATCTCGGCT	1169		
Db	820	TTCTTGAAGTGGTGGCTTAACCTACGGCTACACTAGAGGACAGTATTTGGTATCTCGGCT	879		
QY	1170	CTGCTGAAGCCAGTTACCTTTCGGAAGAGTTGGTAGCTCTTGTATCCGGCAACAACC	1229		
Db	880	CTGCTGAAGCCAGTTACCTTTCGGAAGAGTTGGTAGCTCTTGTATCCGGCAACAACC	939		
QY	1230	ACCGCTGGTAGCGGT	1244		
Db	940	ACCGCTGGTAGCGGT	954		

Search completed: January 17, 2002, 08:54:31
Job time: 7143 sec

THIS PAGE BLANK (USPTO)

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model
Run on: January 17, 2002, 11:50:28 ; Search time 10436.3 seconds
(without alignments)
2856.404 Million cell updates/sec

Title: US-09-242-202A-28
Perfect score: 1807
Sequence: 1 GGTACCTGCCACCATGGCGC.....CTCCACAGCCTCTCCACACA 1807

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 1472140 seqs, 8248589755 residues

Total number of hits satisfying chosen parameters: 2944280

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

GenEmbl :

- 1: gb_ba.*
- 2: gb_htg.*
- 3: gb_in.*
- 4: gb_om.*
- 5: gb_ov.*
- 6: gb_pat.*
- 7: gb_ph.*
- 8: gb_pl.*
- 9: gb_pr.*
- 10: gb_ro.*
- 11: gb_sts.*
- 12: gb_sy.*
- 13: gb_un.*
- 14: gb_vl.*
- 15: em_ba.*
- 16: em_fun.*
- 17: em_hum.*
- 18: em_in.*
- 19: em_om.*
- 20: em_or.*
- 21: em_ov.*
- 22: em_pat.*
- 23: em_ph.*
- 24: em_pl.*
- 25: em_ro.*
- 26: em_sts.*
- 27: em_sy.*
- 28: em_un.*
- 29: em_vl.*
- 30: em_htgo_hum.*
- 31: em_htgo_inv.*
- 32: em_htgo_rod.*
- 33: em_htg_hum.*
- 34: em_htg_inv.*
- 35: em_htg_rod.*
- 36: em_htg_other.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

*

Result No.	Score	Match	Length	DB	ID	Description
1	765	42.3	885	12	SYNPIAN7V	L08875 PIAN7 cloni
2	756.2	41.8	902	12	ATPIVX	V00089 Cloning vec
3	756.2	41.8	902	12	PIVX	X14353 Plasmid pi-
4	756.2	41.8	902	12	SYNPIVXV	L08918 PI-VX cloni
5	755	41.8	2932	6	I07209	I07209 Sequence 3
6	656.2	36.3	3392	12	AB009864	AB009864 Expressio
7	607.4	33.6	7252	6	AX003206	AX003206 Sequence
8	604.4	33.4	5865	6	A60212	A60212 Sequence 8
9	604.4	33.4	5865	6	ARI122288	ARI122288 Sequence
10	604.4	33.4	6028	6	A60209	A60209 Sequence 5
11	604.4	33.4	6028	6	ARI122285	ARI122285 Sequence
12	604.4	33.4	6061	6	A60210	A60210 Sequence 6
13	604.4	33.4	6061	6	ARI122286	ARI122286 Sequence
14	604.4	33.4	6312	6	A60211	A60211 Sequence 7
15	604.4	33.4	6312	6	ARI122287	ARI122287 Sequence
16	604.2	33.4	11795	6	AX027785	AX027785 Sequence
17	604.2	33.4	13254	6	AR038307	AR038307 Sequence
18	604.2	33.4	13254	6	AR038321	AR038321 Sequence
19	604.2	33.4	13254	6	I58596	I58596 Sequence 15
20	604.2	33.4	13254	6	I58610	I58610 Sequence 17
21	603	33.4	1905	6	AR027070	AR027070 Sequence
22	603	33.4	1905	6	I86203	I86203 Sequence 9
23	603	33.4	1968	12	AF087567	AF087567 Cloning v
24	603	33.4	1969	12	AF086840	AF086840 Cloning v
25	603	33.4	1969	12	AF086841	AF086841 Cloning v
26	603	33.4	1969	12	AF086842	AF086842 Cloning v
27	603	33.4	1969	12	AF086843	AF086843 Cloning v
28	603	33.4	1969	12	AF086844	AF086844 Cloning v
29	603	33.4	1969	12	AF086845	AF086845 Cloning v
30	603	33.4	1969	12	AF086846	AF086846 Cloning v
31	603	33.4	1969	12	AF086847	AF086847 Cloning v
32	603	33.4	1969	12	AF086848	AF086848 Cloning v
33	603	33.4	1969	12	AF086849	AF086849 Cloning v
34	603	33.4	1969	12	AF086850	AF086850 Cloning v
35	603	33.4	1969	12	AF086851	AF086851 Cloning v
36	603	33.4	1969	12	AF086852	AF086852 Cloning v
37	603	33.4	1969	12	AF086853	AF086853 Cloning v
38	603	33.4	1969	12	AF087562	AF087562 Cloning v
39	603	33.4	1969	12	AF087563	AF087563 Cloning v
40	603	33.4	1969	12	AF087564	AF087564 Cloning v
41	603	33.4	1969	12	AF087565	AF087565 Cloning v
42	603	33.4	1969	12	AF087566	AF087566 Cloning v
43	603	33.4	1969	12	AF087564	AF087564 Cloning v
44	603	33.4	1969	12	AF087565	AF087565 Cloning v
45	603	33.4	1969	12	AF087566	AF087566 Cloning v

ALIGNMENTS

RESULT 1

SYNPIAN7V

LOCUS

DEFINITION

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

SYNPIAN7V 885 bp DNA circular SYN 26-JUL-1993

PIAN7 cloning vector.

L08875

L08875.1 GI:310776

Synthetic construct DNA.

synthetic construct

artificial sequence.

1 (bases 1 to 885)

Gilbert,W.

Obtained from VecBase 3.0

Unpublished (1991)

These data and their annotation were supplied to GenBank by Will

Gilbert under the auspices of the GenBank Curator Program. PIAN7 -

Cloning vector

ENTRY PIAN7

TITLE PIAN7 - Cloning vector

DATE 12-SEP-1986

#TYPE DNA CIRCULAR

```

#sequence 16-DEC-1986
ACCESSION VB0066
SOURCE artificial
REFERENCE
#number 1
#citation sequence information from New England Biolabs COMMENT
Obtained 12-SEP-1986 from New England Biolabs
by magnetic tape
Revised 16-DEC-1986 by F. Pfeiffer:
449/450 'AT' to 'TA' to match revised sequence of PBR322 COMMENT
PIAN7 is thought to replace pIVX.
COMMENT
The polylinker of PIAN7 contains additional BglII and XbaI
sites within the M13mp8/pUC8 polylinker.
KEYWORDS
CROSSREFERENCE
#parent
VecBase(3):pIVX
#parent
VecBase(3):PBR322, VecSource(3):Poly8, GenBank(50):Ecofgy
PARENT
Features of PIAN7 (885 bp)
residue source
1- 202 1 (c) Tyr-tRNA synthetic (GenBank(50):EcoTgy)
198- 225 1- 28 part 1 of pUC8/M13mp8-polylinker
234- 239 31- 36 part 2 of pUC8/M13mp8-polylinker
240- 880 2522-3162 pBR322
Conflict (cfl) and Mutations (mut): none
FEATURE
POLYLINKER EcoRI-SmaI-BamHI-SalI-PstI-BglII-XbaI-HindIII SELECTION
#suppressor Supf
SUMMARY PIAN7 #length 885 #checksum 9656.
Location/Qualifiers
1..885
/organism="synthetic construct"
/db_xref="taxon:32630"
BASE COUNT 193 a 239 c 236 g 217 t
ORIGIN

Query Match 42.3%; Score 765; DB 12; Length 885;
Best Local Similarity 95.9%; Pred. No. 2.6e-211;
Matches 805; Conservative 0; Mismatches 10; Indels 24; Gaps 1;

QY 478 TTTCGGACTTTTCAAGTGTGCTGGGGAAGATTGCAACCTTCGAAGTCGATGAC 537
DB 1 TTTCGGACTTTTGAAGTGTGCTGGGGAAGATTGCGAACCCTTCGAAGTCGATGAC 60

QY 538 GCAGATTTAGACTGTCTCCCTTTGGCGCTCGGGAACCCACACGGGTAACTGTTT 597
DB 61 GCGAGATTTAGACTGTCTCCCTTTGGCGCTCGGGAACCCACACGGGTAACTGTTT 120

QY 598 ACTGGCTGCTCCCTTATCGGGAAGCGGCGCATCATATCAATACGCGCGCTGTAA 657
DB 121 ACTGGCTGCTCCCTTATCGGGAAGCGGCGCATCATATCAATACGCGCGCTGTAA 180

QY 658 AGTGTTACGTTGAGAAGATTC-----CTGACCCCGCCG 693
DB 181 AGTGTTACGTTGAGAAGATTCCTCCGGGATCGTCGACCTGCAGATCTCTAGAAGCTG 240

QY 694 CGTGTGGGGTTTTTCCATAGCTCGCCCTTCGAGAGATCACAATAATCGAGCT 753
DB 241 CGTGTGGGGTTTTTCCATAGCTCGCCCTTCGAGAGATCACAATAATCGAGCT 300

QY 754 CAAGTCAGAGTGGCGAAACCCGACAGGACTATAAGATACACAGGGGTTTCCCTCGAA 813
DB 301 CAAGTCAGAGTGGCGAAACCCGACAGGACTATAAGATACACAGGGGTTTCCCTCGAA 360

QY 814 GCTCCCTCGTGGCTCTCTCTGTTCCGACCTCGCGCTTACCGGATACCTGTCGCGCTTC 873
DB 361 GCTCCCTCGTGGCTCTCTCTGTTCCGACCTCGCGCTTACCGGATACCTGTCGCGCTTC 420

QY 874 TCCCTTCGGGAACGGTGGCGCTTCTCAATGCTCAGCTCAGCTAGTATCTCAGTTCGTTG 933

```

```

DB 421 TCCTTCGGGAAGCGTGGCGCTTTCTCATAGCTCAGCTGTAGGTATCTCAGTTCGGTGT 480
QY 934 AGGTCGTTCCGCTCAAGCTGGGCTGTGTGCAGAACCCCGGTTTCAGCCCGCTCGG 993
DB 481 AGGTCGTTCCGCTCAAGCTGGGCTGTGTGCAGAACCCCGGTTTCAGCCCGCTCGG 540
QY 994 CCTATCCGGTAACCTATCTTGTAGTCCAAACCCGGTAAGACACGACTTATCGCCACTGG 1053
DB 541 CCTATCCGGTAACCTATCTTGTAGTCCAAACCCGGTAAGACACGACTTATCGCCACTGG 600
QY 1054 CAGCAGCCACTGGTAACAGGATTACGAGCGAGGTATAGGCGGTCTACAGAGTTCT 1113
DB 601 CAGCAGCCACTGGTAACAGGATTACGAGCGAGGTATAGGCGGTCTACAGAGTTCT 660
QY 1114 TGAAGTGTGGCTAACTACGCTACACTAGAACGACAGTATTTGGTATCTGCGCTCTGC 1173
DB 661 TGAAGTGTGGCTAACTACGCTACACTAGAACGACAGTATTTGGTATCTGCGCTCTGC 720
QY 1174 TGAAGTGTGGCTAACTACGCTAAAGAGTTGGTAGCTTTGATCCGCGCAACAAACACCG 1233
DB 721 TGAAGTGTGGCTAACTACGCTAAAGAGTTGGTAGCTTTGATCCGCGCAACAAACACCG 780
QY 1234 CTGGTAGCGGTGTTTTTTTTCGCAAGCAGCAGATTACGCGCAAGAAAAAGGATCT 1292
DB 781 CTGGTAGCGGTGTTTTTTTTCGCAAGCAGCAGATTACGCGCAAGAAAAAGGATCT 839

RESULT 2
ATPIVX 902 bp DNA SYN 28-SEP-1992
LOCUS Cloning vector pi-VX used for screening bacteriophage lambda gene
DEFINITION libraries for specific DNA sequences in Escherichia coli. Probe
sequences are inserted in the vector pi-VX. Contains a polylinker,
and origin of replication (derived from pMB1) and a tyrosine
amber-suppressor gene (synthetic supf gene).
ACCESSION V00089
VERSION V00089.1 GI:58159
KEYWORDS cloning vector; origin of replication; plasmid; polylinker;
suppressor gene.
SOURCE Cloning vector pi-VX.
ORGANISM Cloning vector pi-VX
artificial sequence; vectors.
REFERENCE 1 (bases 1 to 902)
AUTHORS Maniatis,T., Fritsch,E.F. and Sambrook,J.
TITLE Molecular cloning
JOURNAL (in) Maniatis,T., Fritsch,E.F. and Sambrook,J. (Eds.):
MOLECULAR CLONING: 353-361.
Cold Spring Harbor Laboratory (1982)
REFERENCE 2 (bases 1 to 902)
AUTHORS Fitzwater,T., Yang,X.L., Zhang,X.Y. and Polisky,B.
TITLE Mutations affecting RNA-DNA hybrid formation of the Cole1
replication primer RNA. Restoration of RNA I sensitivity to a
copy-number mutant by second-site mutations
JOURNAL J. Mol. Biol. 226 (4), 997-1008 (1992)
MEDLINE 92389346
FEATURES Location/Qualifiers
source
1..902
/organism="Cloning vector pi-VX"
/db_xref="taxon:31856"
rep_origin 600
/note="origin of replication"
BASE COUNT 192 a 250 c 237 g 223 t
ORIGIN

Query Match 41.8%; Score 756.2; DB 12; Length 902;
Best Local Similarity 98.1%; Pred. No. 9.5e-209;
Matches 790; Conservative 0; Mismatches 3; Indels 12; Gaps 2;

QY 471 CGAATTCCTTCGACCTTTTGAAGTGTGTTGGGGGAAGGATTCGAACCTTCAAGT 530
DB 109 CGAATTCCTTCGACCTTTTGAAGTGTGTTGGGGGAAGGATTCGAACCTTCAAGT 168

```

QY	531	CGATGACGGCAGATTTAGAGTCTGCTCCCTTTGSCGCTCGGAACCCACCGGGTAA	590
Db	169	CGATGACGGCAGATTTAGAGTCTGCTCCCTTTGSCGCTCGGAACCCACCGGGTAA	228
QY	591	TGCTTTTACTGGCTGCTCCCTTTATCGGGAAGGGGCGCATCATATCAAAATGACGCGCC	650
Db	229	TGCTTTTACTGGCTGCTCCCTTTATCGGGAAGGGGCGCATCATATCAAAATGACGCGCC	288
QY	651	GCCTGTAAGTGTACGTTGAGAAGAAATTCCTGCAGCCGCCGCTGCTGGCGTTCCTTC	710
Db	289	GCCTGTAAGTGTACGTTGAGAAGAAATTCCTGCAGCCGCCGCTGCTGGCGTTCCTTC	337
QY	711	CATAGGCTCGCCCGCTGACGAGCATCAAAAAATGACGCTCAAGTCAGAGTGGCGA	770
Db	338	CATAGGCTCGCCCGCTGACGAGCATCAAAAAATGACGCTCAAGTCAGAGTGGCGA	397
QY	771	AACCCGACAGGACTATAAAGATACCAAGCGCTTCCCGCTGGAAGCTCCCTCGCGCTCT	830
Db	398	AACCCGACAGGACTATAAAGATACCAAGCGCTTCCCGCTGGAAGCTCCCTCGCGCTCT	457
QY	831	CTGTGTCGACCTCGCGCTTACGGATACCTGTCCGCTTCTCCCTTCGGGAAGCGTG	890
Db	458	CTGTGTCGACCTCGCGCTTACGGATACCTGTCCGCTTCTCCCTTCGGGAAGCGTG	517
QY	891	CGCTTTCTCAATGCTCACGCTGTAGGTATCTCAGTTTCGGTGTAGGTTCGCTCCAAG	950
Db	518	CGCTTTCTCATAGCTCACGCTGTAGGTATCTCAGTTTCGGTGTAGGTTCGCTCCAAG	577
QY	951	CTGGCTGTGTGACGACGAGTGTAGGCGGTGTACAGAGTCTTGAAGTGGTGGCTTAAC	1010
Db	578	CTGGCTGTGTGACGACGAGTGTAGGCGGTGTACAGAGTCTTGAAGTGGTGGCTTAAC	697
QY	1011	CGCTTACGACTAGAGGACAGTATTTGGTATCTGCGCTCTGCTGGAAGCGATTAAC	1070
Db	638	CGCTTACGACTAGAGGACAGTATTTGGTATCTGCGCTCTGCTGGAAGCGATTAAC	697
QY	1071	AGGATTAGCAGGAGGATGTAGGCGGTGTACAGAGTCTTGAAGTGGTGGCTTAAC	1130
Db	698	AGGATTAGCAGGAGGATGTAGGCGGTGTACAGAGTCTTGAAGTGGTGGCTTAAC	757
QY	1131	TACGGCTACACTAGAAGGACAGTATTTGGTATCTGCGCTCTGCTGGAAGCGATTAAC	1190
Db	758	TACGGCTACACTAGAAGGACAGTATTTGGTATCTGCGCTCTGCTGGAAGCGATTAAC	816
QY	1191	GGAAAAGAGTGTGAGTCTTGTATCGGCAACAAACACCGCTGCTAGCGGTGTTT	1250
Db	817	GGAAAAGAGTGTGAGTCTTGTATCGGCAACAAACACCGCTGCTAGCGGTGTTT	876
QY	1251	TTTGTGTAAGCAGGATTAACGC	1275
Db	877	TTTGTGTAAGCAGGATTAACGC	901
RESULT	3		
LOCUS	PPIVX		
DEFINITION	PPIVX	902 bp	DNA
ACCESSION	X14353		
VERSION	X14353.1	GI:58268	
KEYWORDS	artificial sequence; origin of replication; plasmid; polylinker; suppressor transfer RNA.		
SOURCE	synthetic construct.		
ORGANISM	Plasmid synthetic construct		
REFERENCE	1 (bases 1 to 902)		
AUTHORS	Seed,B.		
TITLE	Purification of genomic sequences from bacteriophage libraries by recombination and selection in vivo		
JOURNAL	Nucleic Acids Res. 11 (8), 2427-2445 (1983)		
MEDLINE	83220795		
FEATURES	Location/Qualifiers		
source	1. .902		

/organism="synthetic construct"
 /plasmid="pi-VX"
 /db_xref="taxon:32630"
 misc_feature 1. .109
 /note="polylinker"
 trna 110. .212
 /note="suppressor trna"
 misc_feature 213. .902
 /note="origin of replication"
 BASE COUNT 192 a 249 c 238 g 223 t
 ORIGIN

Query Match	41.8%;	Score 756.2;	DB 12;	Length 902;
Best Local Similarity	98.1%;	Pred. No. 9.5e-209;		
Matches 790;	Conservative 0;	Mismatches 3;	Indels 12;	Gaps 2;
QY	471	CGAATCTTTTCGGACTTTTGAAGTGTATGTTGGGGGAAGGATTCGAACCTTCGAAGT	530	
Db	109	CGAATCTTTTCGGACTTTTGAAGTGTATGTTGGGGGAAGGATTCGAACCTTCGAAGT	168	
QY	531	CGATGACGGCAGATTTAGAGTCTGCTCCCTTTGGCCGCTCGGGAACCCACCGGGTAA	590	
Db	169	CGATGACGGCAGATTTAGAGTCTGCTCCCTTTGGCCGCTCGGGAACCCACCGGGTAA	228	
QY	591	TGCTTTTACTGGCTGCTCCCTTTATCGGGAAGGGGCGCATCATATCAAAATGACGCGCC	650	
Db	229	TGCTTTTACTGGCTGCTCCCTTTATCGGGAAGGGGCGCATCATATCAAAATGACGCGCC	288	
QY	651	GCTGTAAGTGTACGTTGAGAAGAAATTCCTGCAGCCGCCGCTGCTGGCGTTCCTTC	710	
Db	289	GCTGTAAGTGTACGTTGAGAAGAAATTCCTGCAGCCGCCGCTGCTGGCGTTCCTTC	337	
QY	711	CATAGGCTCGCCCGCTGACGAGCATCAAAAAATGACGCTCAAGTCAGAGTGGCGA	770	
Db	338	CATAGGCTCGCCCGCTGACGAGCATCAAAAAATGACGCTCAAGTCAGAGTGGCGA	397	
QY	771	AACCCGACAGGACTATAAAGATACCAAGCGCTTCCCGCTGGAAGCTCCCTCGCGCTCT	830	
Db	398	AACCCGACAGGACTATAAAGATACCAAGCGCTTCCCGCTGGAAGCTCCCTCGCGCTCT	457	
QY	831	CTGTGTCGACCTCGCGCTTACGGATACCTGTCCGCTTCTCCCTTCGGGAAGCGTG	890	
Db	458	CTGTGTCGACCTCGCGCTTACGGATACCTGTCCGCTTCTCCCTTCGGGAAGCGTG	517	
QY	891	CGCTTTCTCAATGCTCACGCTGTAGGTATCTCAGTTTCGGTGTAGGTTCGCTCCAAG	950	
Db	518	CGCTTTCTCATAGCTCACGCTGTAGGTATCTCAGTTTCGGTGTAGGTTCGCTCCAAG	577	
QY	951	CTGGCTGTGTGACGACGAGTGTAGGCGGTGTACAGAGTCTTGAAGTGGTGGCTTAAC	1010	
Db	578	CTGGCTGTGTGACGACGAGTGTAGGCGGTGTACAGAGTCTTGAAGTGGTGGCTTAAC	637	
QY	1011	CGCTTACGACTAGAGGACAGTATTTGGTATCTGCGCTCTGCTGGAAGCGATTAAC	1070	
Db	638	CGCTTACGACTAGAGGACAGTATTTGGTATCTGCGCTCTGCTGGAAGCGATTAAC	697	
QY	1071	AGGATTAGCAGGAGGATGTAGGCGGTGTACAGAGTCTTGAAGTGGTGGCTTAAC	1130	
Db	698	AGGATTAGCAGGAGGATGTAGGCGGTGTACAGAGTCTTGAAGTGGTGGCTTAAC	757	
QY	1131	TACGGCTACACTAGAAGGACAGTATTTGGTATCTGCGCTCTGCTGGAAGCGATTAAC	1190	
Db	758	TACGGCTACACTAGAAGGACAGTATTTGGTATCTGCGCTCTGCTGGAAGCGATTAAC	816	
QY	1191	GGAAAAGAGTGTGAGTCTTGTATCGGCAACAAACACCGCTGCTAGCGGTGTTT	1250	
Db	817	GGAAAAGAGTGTGAGTCTTGTATCGGCAACAAACACCGCTGCTAGCGGTGTTT	876	
QY	1251	TTTGTGTAAGCAGGATTAACGC	1275	
Db	877	TTTGTGTAAGCAGGATTAACGC	901	

```

RESULT 4
SYNPIVXV 902 bp DNA circular SYN 26-JUL-1993
LOCUS pi-VX cloning vector used for screening bacteriophage lambda.
DEFINITION L08918
ACCESSION L08918.1 GI:310786
VERSION 1
KEYWORDS Synthetic construct DNA.
SOURCE synthetic construct
ORGANISM artificial sequence.
REFERENCE 1 (bases 1 to 902)
AUTHORS Gilbert, W.
TITLE Obtained from VecBase 3.0
JOURNAL Unpublished (1991)
COMMENT These data and their annotation were supplied to GenBank by Will
Gilbert under the auspices of the GenBank Curator Program. pi-VX -
Cloning vector used for screening bacteriophage lambda ENTRY PIVX
#TYPE DNA CIRCULAR

TITLE pi-VX - Cloning vector used for screening bacteriophage
lambda
DATE 17-DEC-1982
ACCESSION V00010
SOURCE artificial
COLLECTION ATCC 39083
REFERENCE
#number 1
#authors Seed B.
#journal Nucl. Acids Res. (1983) 11: 2427-2445
#title Purification of genomic sequences from bacteriophage libraries
and selection in vivo.
#comment bases 1-902
#number 2
#authors Maniatis T., Fritsch E.F., Sambrook J.
#book (IN) Maniatis T., Fritsch E.F., Sambrook J.;
MOLECULAR CLONING: 353-361;
COLD SPRING HARBOR LABORATORY (1982).
#comment BASES 1 TO 902
REFERENCE 3
#authors Backman K., Betlach M., Boyer H.W., Yanofsky S.
#journal Cold Spring Harb Symp Quant Biol 43, 69-76 (1980)
#title Genetic and Physical Studies on the replication of Cole1-type
plasmids
#comment the sequence of the piVX origin is from pBR345 as published in
this paper
COMMENT
The origin of replication is derived from pMB1 via pBR345 and
sequence given for pBR345 differs at one position from that of
pBR322. This
may be a sequencing error.
COMMENT
The ATCC strain 39083 contains piVX in combination with plasmid
p3.
COMMENT
from GenBank
USED FOR SCREENING BACTERIOPHAGE LAMBDA GENE LIBRARIES FOR
SPECIFIC DNA SEQUENCES IN ESCHERICHIA COLI. PROBE SEQUENCES
ARE INSERTED IN THE VECTOR PI-VX. CONTAINS A POLYLINKER, AND
ORIGIN OF REPLICATION AND A TYROSINE
AMBER-SUPPRESSOR GENE (SYNTHETIC SUPP GENE).
KEY FROM TO
ORGRPL 500 600 ORIGIN OF REPLICATION KEYWORDS
suppressor gene \
polylinker \
origin of replication
CROSSREFERENCE
#identical
GenBank(50):PIVX, EMBL(11):ATPIVX
#parent

VecBase(3):pBR322, GenBank(50):EcoTgy, VecSource(3):OriMB1
#offspring
VecBase(3):pIAN7, VecBase(3):mWB2344
PARENT
Features pf piVX (902 bp)
residue source
4- 37 pBR322
1- 115 piVX-Polylinker
111- 317 207- 1 (c) Tvr-trNA synthetic (GenBank(50):EcoTgy)
320- 901 1- 583 OriMB1
320- 901 2522-3104 pBR322
Conflict (cfl) and Mutations (mut):
piVX source
cfl 811-12 T.C TAC 492-94 OriMB1
cfl 811-12 T.C TAC 3013-15 pBR322
FEATURE
SUPP POLYLINKER EcoRI-ClaI-HindIII-XbaI-BglII-PstI-BamHI-EcoRI
SELECTION
#suppressor SupF
SUMMARY Location/Qualifiers
1. 902
/organism="synthetic construct"
/db_xref="taxon:32630" 223 t
BASE COUNT 192 a 250 c 237 g
ORIGIN
Query Match 41.8%; Score 756.2; DB 12; Length 902;
Best Local Similarity 98.1%; Pred. No. 9.5e-209; Indels 12; Gaps 2;
Matches 790; Conservative 0; Mismatches 3;

QY 471 CGAATCTTTTCGGACCTTTGAAAGTGATGGTGGGGGAAGGATTCGAACCTTCGAAGT 530
Db 109 CGAATCTTTTCGGACCTTTGAAAGTGATGGTGGGGGAAGGATTCGAACCTTCGAAGT 168
QY 531 CGATGACGCGAGATTAGAGTCTGCTCCCTTTGGCGCTCGGGAAACCCACCACGCGTAA 590
Db 169 CGATGACGCGAGATTAGAGTCTGCTCCCTTTGGCGCTCGGGAAACCCACCACGCGTAA 228
QY 591 TCGCTTTTACTGGCTGCTCCCTTTATCGGAAGCGGGCGCATCATATCAAAATGACGCGCC 650
Db 229 TCGCTTTTACTGGCTGCTCCCTTTATCGGAAGCGGGCGCATCATATCAAAATGACGCGCC 288
QY 651 GCTGTAAAGTGTACGTTGAGAAAGAAATTCCTGCACGCCCGCGCTTCTGCGGCTTTTC 710
Db 289 GCTGTAAAGTGTACGTTGAGAAAGAAATTCCTGCACGCCCGCGCTTCTGCGGCTTTTC 337
QY 711 CATAGGCTCGGCCCTTCGACGAGCATCAAAAATCGAGCTCAAGTCAGAGTGGCGA 770
Db 338 CATAGGCTCGGCCCTTCGACGAGCATCAAAAATCGAGCTCAAGTCAGAGTGGCGA 397
QY 771 AACCCGACAGGACTATAAGATACCAAGGCGTTTCCCTCGGAAAGCTCCCTTCGCGCTCT 830
Db 398 AACCCGACAGGACTATAAGATACCAAGGCGTTTCCCTCGGAAAGCTCCCTTCGCGCTCT 457
QY 831 CCGTTCGACCGCTCGCGCTTACCGGATACCTGTCGCCCTTCTCCCTTCGGGAAGCGTG 890
Db 458 CCGTTCGACCGCTCGCGCTTACCGGATACCTGTCGCCCTTCTCCCTTCGGGAAGCGTG 517
QY 891 GCCTTTCTCAAGTCTACGCTGACGCTGTAGTATCTCAGTTCGGGTAGGTGCTCGTCCAAG 950
Db 518 GCCTTTCTCATAGCTACGCTGACGCTGTAGTATCTCAGTTCGGGTAGGTGCTCGTCCAAG 577
QY 951 CTGGGCTGTGTGCACGAACCCCGCTTACGCCCGACCGCTTACGCCCTTATCCGGTAACTAT 1010
Db 578 CTGGGCTGTGTGCACGAACCCCGCTTACGCCCGACCGCTTACGCCCTTATCCGGTAACTAT 637
QY 1011 CGCTTTGAGTCCAAACCGGTAAGACAGCACTTATCGCACTGGCAGCAGCACTGGTAAC 1070
Db 638 CGTCTTGAGTCCAAACCGGTAAGACAGCACTTATCGCACTGGCAGCAGCACTGGTAAC 697

```



```
Qy 1071 AGGATTAGCAGAGGAGGTATGTAGCGGTGCTACAGAGTTCCTTGAAGTGGTGGCCCTAAC 1130
|||||
Db 698 AGGATTAGCAGAGGAGGTATGTAGCGGTGCTACAGAGTTCCTTGAAGTGGTGGCCCTAAC 757
|||||
Qy 1131 TAGCGCTACACTAGAAGGACAGTATTTGGTATCTGCGCTCTGCTGAAGCAGATTACCTTC 1190
|||||
Db 758 TAGCGCTACACTAGAAGGACAGTATTTGGTATCTGCGCTCTGCTGAAGCAGATT-CCTTC 816
|||||
Qy 1191 GGAAGAGAGTTGGTAGCTCTTGATCCGGCAAAACAAACACCGCTGGTAGCGGTGGTTTTT 1250
|||||
Db 817 GGAAGAGAGTTGGTAGCTCTTGATCCGGCAAAACAAACACCGCTGGTAGCGGTGGTTTTT 876
|||||
Qy 1251 TTGTTTGCAGCAGCAGATTACGC 1275
|||||
Db 877 TTTGTTTGCAGCAGCAGATTACGC 901
|||||

RESULT 5
107209/c 107209 2932 bp PAT 02-DEC-1994
LOCUS Sequence 3 from Patent EP 0330191.
DEFINITION
ACCESSION 107209
VERSION 107209.1 GI:590060
KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.
REFERENCE
1 (bases 1 to 2932)
Seed,B.D.,D.O.M.B., Allen,J., Aruffo,A., Camerini,D., Lauffer,L.D.,
Oquendo,C.P., Simmons,D., Stamenkovic,I. and Stengelin,S.D.
Rapid immunoselection cloning method
JOURNAL Patent: EP 0330191-A2 3 30-AUG-1989;
FEATURES
Location/Qualifiers
source 1..2932
BASE COUNT 774 a 699 c 682 g 777 t
ORIGIN

Query Match 41.88; Score 755; DB 6; Length 2932;
Best Local Similarity 97.68; Pred. No. 2.2e-208;
Matches 788; Conservative 0; Mismatches 15; Indels 4; Gaps 2;

Qy 469 ATCGAATTCCTTCGGACTTTTGAAGTAGTGTTGGGGGAAGGATTGCAACCTTCGAA 528
|||||
Db 804 AGCAGATTCCTTCGGACTTTTGAAGTAGTGTTGGGGGAAGGATTGCAACCTTCGAA 745
|||||
Qy 529 GTCGATGACGCGCAGATTAGAGTCTGCTCCCTTTGGCGCTCGGGAACCCACCACCGGTT 588
|||||
Db 744 GTCGATGACGCGCAGATTAGAGTCTGCTCCCTTTGGCGCTCGGGAACCCACCACCGGTT 685
|||||
Qy 589 AATGCTTTTACTGCGCTGCTCCCTTATCGGAAGCGGGGGGCATCATATCAAAATGACGCG 648
|||||
Db 684 AATGCTTTTACTGCGCTGCTCCCTTATCGGAAGCGGGGGGCATCATATCAAAATGACGCG 625
|||||
Qy 649 CCGCTGTAAGCTGTACGTTGAGAAAGTAATTCCTGACGCCCGCGCTGCTGCGGCTTTT 708
|||||
Db 624 CCGCTGTAAGCTGTACGTTGAGAAAGA---CCGCGGTAATTCGCGGCTGCTGCGGCTTTT 568
|||||
Qy 709 TCCATAGGCTCCGCGCCCTCGACGAGCATCACAAAATTCGACGCTCAAGTCAGAGGTGGC 768
|||||
Db 567 TCCATAGGCTCCGCGCCCTCGACGAGCATCACAAAATTCGACGCTCAAGTCAGAGGTGGC 508
|||||
Qy 769 GAAACCGCAGGAGCTATAAGATACAGGCGTTTCCCGCTGGAAGCTCCCTCGTGGCGCT 828
|||||
Db 507 GAAACCGCAGGAGCTATAAGATACAGGCGTTTCCCGCTGGAAGCTCCCTCGTGGCGCT 448
|||||
Qy 829 CTCCTGTTCCGACCTCCGCTTACCGATACCTGTCGCGCTTTCTCCCTTCGGGAAGCG 888
|||||
Db 447 CTCCTGTTCCGACCTCCGCTTACCGATACCTGTCGCGCTTTCTCCCTTCGGGAAGCG 388
|||||
Qy 889 TGGCGCTTTCTCAATGCTCAGCTGTAGGTATCTCAGTTGCGTGTAGTTCGTTCCGCTCCA 948
|||||
```

```
Db 387 TGGCGCTTTCTCATAGCTCAGCGTGTAGGTATCTCAGTTGCGTGTAGGTGCTTCGCTCCA 328
Qy 949 AGCTGGGCTGTGTGCAGAACCCCGCTTACGCCGCGCTGCCTTATTCGGGTAAC 1008
|||||
Db 327 AGCTGGGCTGTGTGCAGAACCCCGCTTACGCCGCGCTGCCTTATTCGGGTAAC 268
|||||
Qy 1009 ATCGCTCTTGAGTCCAAACCCGGTAAGACACGACTTATCGCCACTGGCAGCAGCCACTGGTA 1068
|||||
Db 267 ATCGCTCTTGAGTCCAAACCCGGTAAGACACGACTTATCGCCACTGGCAGCAGCCACTGGTA 208
|||||
Qy 1069 ACAGGATTAGCAGACGAGGTATGTAGCGGTGCTACAGAGTTCTTGAAGTGGTGGCCTA 1128
|||||
Db 207 ACAGGATTAGCAGACGAGGTATGTAGCGGTGCTACAGAGTTCTTGAAGTGGTGGCCTA 148
|||||
Qy 1129 ACTACGGCTACACTAGAAGGACAGTATTTGGTATCTGCGCTCTGCTGAAGCAGCTTACT 1188
|||||
Db 147 ACTACGGCTACACTAGAAGGACAGTATTTGGTATCTGCGCTCTGCTGAAGCAGCTT-CCT 89
|||||
Qy 1189 TCGGAAAAAGAGTTGGTAGCTCTTTGATCCGGCAAAACAAACACCGCTGGTAGCGGTGGTT 1248
|||||
Db 88 TCGGAAAAAGAGTTGGTAGCTCTTTGATCCGGCAAAACAAACACCGCTGGTAGCGGTGGTT 29
|||||
Qy 1249 TTTTGTGTTGCAACGACAGATTACGC 1275
|||||
Db 28 TTTTGTGTTGCAACGACAGATTACGC 2

RESULT 6
AB009864
LOCUS 3392 bp DNA circular SYN 26-DEC-1997
DEFINITION Expression vector pME18S-FL3, complete sequence.
ACCESSION AB009864
VERSION AB009864.1 GI:2723416
KEYWORDS
SOURCE expression vectors plasmid:pME18S-FL3 DNA.
ORGANISM Plasmid vectors
REFERENCE 1 (bases 1 to 3392)
AUTHORS Maruyama,K. and Sugano,S.
TITLE Direct Submission
JOURNAL Submitted (16-DEC-1997) to the DDBJ/EMBL/GenBank databases. Sumio
Sugano, The Institute of Medical Science, University of Tokyo,
Department of Virology; 4-6-1, Shirokanedai, Minato-ku, Tokyo 108,
Japan (E-mail: ssugano@ims.u-tokyo.ac.jp, Tel:81-3-5449-5286,
Fax:81-3-5449-5416)
REFERENCE 2 (bases 1 to 3392)
AUTHORS Maruyama,K. and Sugano,S.
TITLE pME18S-FL3: a versatile expression vector
JOURNAL Published Only in DataBase (1997) In press
FEATURES
Location/Qualifiers
source 1..3392
/organism="vectors"
/promoter 1..638
/polyA_site 1229..1427
BASE COUNT 810 a 902 c 818 g 862 t
ORIGIN

Query Match 36.38; Score 656.2; DB 12; Length 3392;
Best Local Similarity 70.88; Pred. No. 1.3e-179;
Matches 990; Conservative 0; Mismatches 288; Indels 121; Gaps 4;

Qy 15 TGGCGCGGATTCCTTATCAGTGAAGTTGGTGACATATTTATGTTTATCAGTGATAAG 74
|||||
Db 828 TGGCGCGGATTCCTTATCAGTGAAGTTGGTGACATATTTATGTTTATCAGTGATAAG 887
|||||
Qy 75 TGTCAGCATGACAAAGTTGCAGCCGAATACAGTATCCGTGCGCGCCCTGGAGCTGTGA 134
|||||
Db 888 TGTCAGCATGACAAAGTTGCAGCCGAATACAGTATCCGTGCGCGCCCTAGACCTGTGA 947
|||||
Qy 135 ACGAGGTGCGGTAGACGGTCTGACGACACCGCAAACTGCGCGGAACGGTTGGGGGTGCAGC 194
|||||
```

Db 948 ACAGGTCGCGTAGACGGTCTACGACACGCAAACTGGCGGAACGGTTGGGGTTTCAGC 1007
QY 195 AGCGCGCGCTTTACTGCGACTTCAGGAACAGCGGGCGCTTAAAGGGCCATATGTTGAGT 254
Db 1008 AGCGCGCGCTTTACTGCGACTTCAGGAACAGCGGGCGCTGCTCGACGCACTGGCCGAAG 1067
QY 255 GGATGC-----CCTTGACCCAGCGGGGATGGGGGAGACC 289
Db 1068 CCATGCTGGCGGAGAAATCATAGACATTCGGTGGCGAGAGCGGACGAGCTGGCGCTCAT 1127
QY 290 TGTAGTCAGAGCCCCCGGGCAGCAGCAATGCCGCTCTTCCCTGCGAGTAGTGA 349
Db 1128 TTCTGACTGGGAATCCCGCAGCTTCAGGCAGCGCTGCTCGCCTACCGCCACACAATG 1187
QY 350 GTGAGTCCCTCTCTGCGCCCTGGAAGTTGCCACTCCAGTCCAGTCCACCAGCCTTGTCT-- 406
Db 1188 GCACCCCGTCTCGAGCTGCGAGTCCGCGCGCTAGACTAGCTAGAGAAAAACCTCC 1247
QY 407 -----AATAAAATTAAGTTGCATCATTTTGTCTGACTAGTGT 444
Db 1248 ACACCTCCCTGNACTGAAACATAAATGAATGCAATTTGTTGTTAACTTGTTAT 1307
QY 445 CCTCTAATATTAAGCTTGATATCGAATCTTTGCGGACTTTTGAAGTGAATGTTGT 504
Db 1308 TGCAGCTTATAATGTTTACAAATAAGCAATAGCATCACAAATTTTCAAAATAAAGCAAT 1367
QY 505 GGGGAGGATTCGAACCTTCGAAGTCGATGAGCGGAGATTTAGAGTCTCTCCCTTTGG 564
Db 1368 TTTTTCACGATTCATGATGTTGTGTTGTTGCCAACTCATCAATGATCTATCATGTCGT 1427
QY 565 CCGCTCGGGAACCCACACGGGTAAATGCTTTTACTGGCTGTCTCCCTTATCGGGAAGCG 624
Db 1428 GATCCCGGGTACCGAGCTCGAATTAATCTCTTCGCTCTCTCGCTCTACTGACTGCT 1487
QY 625 GGGCGCATCATATAAANFAGCGG----- 648
Db 1488 GCGCTCGGTGCTCGGCTGGCGGAGCGGGTATCAGCTCACTCAAGGCGGTAAATACGGTT 1547
QY 649 -----CCGCTGTAAAGTGTACGTTGAGAAAGAAATTCCTGACGCC 688
Db 1548 ATCCACAGAAATCAGGGATACGCGAGGAAGAACATGTGAGCAAAAGCGCCAAAGGC 1607
QY 689 C-----GCGCGTTGCTGGCGTTTTCATAGGCTCCGCCCCCTGACGA 733
Db 1608 CAGGAACGTAAGGCGGTTGCTGGCGTTTTCATAGGCTCCGCCCCCTGACGA 1667
QY 734 GCATCAAAAATCGAGCTCAAGTCAGAGGTGCGGAACCCGACAGGACTATAAAGATA 793
Db 1668 GCATCAAAAATCGAGCTCAAGTCAGAGGTGCGGAACCCGACAGGACTATAAAGATA 1727
QY 794 CCAGCGTTTCCCTCTGGAAGCTCCCTGCTGCTCTCTGCTTCCGACCTTCCGCTTAC 853
Db 1728 CCAGCGTTTCCCTCTGGAAGCTCCCTGCTGCTCTCTGCTTCCGACCTTCCGCTTAC 1787
QY 854 CGGATACCTGTCGCGCTTCTCCCTTCGGGAAGCGTGGCGCTTCTCAATGCTCACGCTG 913
Db 1788 CGGATACCTGTCGCGCTTCTCCCTTCGGGAAGCGTGGCGCTTCTCAATGCTCACGCTG 1847
QY 914 TAGGTATCTAGTTCGGTGTAGTGTTCGCTCAAGTGGGCTGTGTGACGAACCCCC 973
Db 1848 TAGGTATCTAGTTCGGTGTAGTGTTCGCTCAAGTGGGCTGTGTGACGAACCCCC 1907
QY 974 GGTTCAGCCGACCGCTGCGCTTATCCGGTAACATATCTGCTTCTGAGTCCAAACCCGGTAAG 1033
Db 1908 GTTTCAGCCGACCGCTGCGCTTATCCGGTAACATATCTGCTTCTGAGTCCAAACCCGGTAAG 1967
QY 1034 ACAGGACTTATCGCACTGGCAGAGCCACTTGTAAACAGGATTTAGCAGAGGAGTATGT 1093
Db 1968 ACAGGACTTATCGCACTGGCAGAGCCACTTGTAAACAGGATTTAGCAGAGGAGTATGT 2027
QY 1094 AGCGGCTGCTACAGAGTCTTGAAGTGGTGGCCCTAACTACGGCTACACTAGGAAGACAGT 1153

Db 2028 AGCGGTGCTACAGAGTCTTCTTGAAGTGGTGGCTAACTACGGCTACACTAGAGGACAGT 2087
QY 1154 ATTGTGTAATCTGCGTCTGCTGAAGCCAGTTACTCTCGGAAAAAGAGTTGGTAGCTCTTG 1213
Db 2088 ATTGTGTAATCTGCGTCTGCTGAAGCCAGTTACTCTCGGAAAAAGAGTTGGTAGCTCTTG 2147
QY 1214 ATCCGCAAAACAAACACCGCTGCTGAGCGGTGTTTGTGTTTGTGCAACGACGAGATTAC 1273
Db 2148 ATCCGCAAAACAAACACCGCTGCTGAGCGGTGTTTGTGTTTGTGCAACGACGAGATTAC 2207
QY 1274 GGCAGAAAAAAGGATCT 1292
Db 2208 GGCAGAAAAAAGGATCT 2226
RESULT 7
AX003206 AX003206 7252 bp DNA PAT 24-AUG-2000
LOCUS Sequence 57 from Patent WO9932646.
DEFINITION AX003206
ACCESSION AX003206
VERSION AX003206.1 GI:9927068
KEYWORDS synthetic construct.
SOURCE synthetic construct.
ORGANISM artificial sequence.
REFERENCE 1 (bases 1 to 7252)
AUTHORS Carroll,M.W. and Mitrophanous,K.
TITLE Equine infectious anaemia virus (eIav) based
JOURNAL Patent: WO 9932646-A 57 01-JUL-1999;
CARROLL MILES WILLIAM (GB); MITROPHANOUS KYRIACOS (GB)
FEATURES
Location/Qualifiers
1..7252
/organism="synthetic construct"
/db_xref="taxon:32630"
/note="complete sequence of pSC65"
BASE COUNT 1942 a 1686 c 1751 g 1873 t
ORIGIN
Query Match 33.6%; Score 607.4; DB 6; Length 7252;
Best Local Similarity 97.5%; Pred. No. 2.1e-165;
Matches 617; Conservative 0; Mismatches 16; Indels 0; Gaps 0;
QY 660 TGTTAGCTTGAAGAAATTCCTGCGAGCCGCGCGTGTGCTGCGTGTTCCTATAGGCTC 719
Db 5062 TATACCGGTGAGAACATAGACTGACTCGGGCGCGTGTGCTGCGTGTTCCTATAGGCTC 5121
QY 720 CGCCCGCTCACGAGCATCAAAAATCGACGCTCAAGTCAGAGGTGGCGAAACCCGACA 779
Db 5122 CGCCCGCTCACGAGCATCAAAAATCGACGCTCAAGTCAGAGGTGGCGAAACCCGACA 5181
QY 780 GGACTATAAAGATACACAGCGTTTCCCGCTTGAAGCTCCCTGCTGCGCTCTCCTGTTCCG 839
Db 5182 GGACTATAAAGATACACAGCGTTTCCCGCTTGAAGCTCCCTGCTGCGCTCTCCTGTTCCG 5241
QY 840 ACCCTGCGCTTACCAGATACCTGTCGCGCTTCTCCCTTCGGGAAGCGTGGCGCTTCT 899
Db 5242 ACCCTGCGCTTACCAGATACCTGTCGCGCTTCTCCCTTCGGGAAGCGTGGCGCTTCT 5301
QY 900 CAATGCTACGCTGTAGTATCTCAGTTCGCTGTAGTTCGCTCGCTCAAGCTGGGCTGT 959
Db 5302 CAATGCTACGCTGTAGTATCTCAGTTCGCTGTAGTTCGCTCGCTCAAGCTGGGCTGT 5361
QY 960 GTGCACGAACCCCGCTTACGCCCGACCGCTGCGCTTATCCGGTAACATATCGCTTTGAG 1019
Db 5362 GTGCACGAACCCCGCTTACGCCCGACCGCTGCGCTTATCCGGTAACATATCGCTTTGAG 5421
QY 1020 TCCAAACCCGTAAGACAGACTTATCCCACTGGCAGACCCACTGCTACAGGATTAGC 1079
Db 5422 TCCAAACCCGTAAGACAGACTTATCCCACTGGCAGACCCACTGCTACAGGATTAGC 5481
QY 1080 AGACGAGGTATGTAGCGGTGCTACAGAGTCTTGAAGTGGTGGCTTAACCTAGCGCTAC 1139

Db 5482 AGAGCAGGATGTAGGCGGTGCTACAGAGTTCTTGAAGTGTGGCCTAACTACGCTAC 5541

QY 1140 ACTAGAAGACAGTATTTGGTATCTCGCTCTGCTGAAGCCAGTTACCTTCGGAAGAA 1199

Db 5542 ACTAGAAGACAGTATTTGGTATCTCGCTCTGCTGAAGCCAGTTACCTTCGGAAGAA 5601

QY 1200 GTTGGTAGCTCTGTAGTCGGCAAAACAAACACCGCTGTGTAGCGGTGGTTTTTTTGG 1259

Db 5602 GTTGGTAGCTCTGTAGTCGGCAAAACAAACACCGCTGTGTAGCGGTGGTTTTTTTGG 5661

QY 1260 AAGCAGCAGATTACGGCGCAGAAAAAAGGATCT 1292

Db 5662 AAGCAGCAGATTACGGCGCAGAAAAAAGGATCT 5694

RESULT 8

A60212 A60212 5865 bp DNA circular PAT 06-MAR-1998

LOCUS Sequence 8 from Patent WO9708330.

DEFINITION A60212

ACCESSION A60212

VERSION A60212.1 GI:3715220

KEYWORDS .

SOURCE unidentified.

ORGANISM unidentified.

REFERENCE 1 (bases 1 to 5865)

AUTHORS Collins,M.K., Weiss,R.A., Takeuchi,Y. and Cosset,F.

TITLE EXPRESSION SYSTEMS

JOURNAL Patent: WO 9708330-A 8 06-MAR-1997;

FEATURES

LOCATION/Qualifiers

1..5865

/organism="unidentified"

/db_xref="taxon:32644"

BASE COUNT 1492 a 1602 c 1390 g 1373 t 8 others

ORIGIN

Query Match 33.4%; Score 604.4; DB 6; Length 5865;

Best Local Similarity 91.4%; Pred. No. 1.6e-164;

Matches 638; Conservative 0; Mismatches 60; Indels 0; Gaps 0;

QY 595 TTACTGGCTGTCTCCCTTATCGGGAAGCGGGCGGCATCATATCAAAATGACGCGCGCTG 654

Db 3752 TTATCATGTCTGGATCCAGATCTGGGCCCATGGCGCGGATCGATNNNNACATGTGAG 3811

QY 655 TAAAGTGTACGTTGAGAAGAAATTCCTGCAGCCCGCGGTGGTGGCGTTTTTCCATA 714

Db 3812 CAAAAGGCCAGCAAAAGGCCAGGAACCGTAAAGAGCCGCTTGTGGCGTTTTTCCATA 3871

QY 715 GCGTCGCCCGCCCTGACGAGCATCACAAAATCGACGCTCAAGTCAGAGGTGGCGAAACC 774

Db 3872 GCGTCGCCCGCCCTGACGAGCATCACAAAATCGACGCTCAAGTCAGAGGTGGCGAAACC 3931

QY 775 CGACAGGACTATAAAGATACCAAGCGGTTTTCCCGCTGGAAGCTCCCTCGCGCTCTCCTG 834

Db 3932 CGACAGGACTATAAAGATACCAAGCGGTTTTCCCGCTGGAAGCTCCCTCGCGCTCTCCTG 3991

QY 835 TTCCGACCTGCGCGTTACGGATACCTGTCGCCCTTTCTCCCTTCGGAAGCGTGGCGC 894

Db 3992 TTCCGACCTGCGCGTTACGGATACCTGTCGCCCTTTCTCCCTTCGGAAGCGTGGCGC 4051

QY 895 TTCTCAATGCTCACGCTGTAGTATCTCAGTTTCGCTGAGGTCGTTCCCTCCAAAGCTGG 954

Db 4052 TTCTCAATGCTCACGCTGTAGTATCTCAGTTTCGCTGAGGTCGTTCCCTCCAAAGCTGG 4111

QY 955 GCTGTGTGCAGCAACCCCGGTTACGCCGACCGCTGCGCTTTATCCGGTAACTATCGTC 1014

Db 4112 GCTGTGTGCAGCAACCCCGGTTACGCCGACCGCTGCGCTTTATCCGGTAACTATCGTC 4171

QY 1015 TTGAGTCCAAACCCGGTAAAGACAGGACTTATCGCCACTGGCAGCAGCAGCTGGTAAACAGGA 1074

Db 4172 TTGAGTCCAAACCCGGTAAAGACAGGACTTATCGCCACTGGCAGCAGCAGCTGGTAAACAGGA 4231

QY 1075 TTAGCAGCGAGGTATGTAGCGGTGCTACAGAGTTCTTGAAGTGTGGCCTAACTACG 1134

Db 4232 TTAGCAGCGAGGTATGTAGCGGTGCTACAGAGTTCTTGAAGTGTGGCCTAACTACG 4291

QY 1135 GCTACACTAGAGGACAGTATTTGGTATCTGCGCTCTGCTGAAGCCAGTTACCTTCGGAA 1194

Db 4292 GCTACACTAGAGGACAGTATTTGGTATCTGCGCTCTGCTGAAGCCAGTTACCTTCGGAA 4351

QY 1195 AAGAGTTGGTAGCTCTGTATCGCGCAAAACAAACACCGCTGTAGCGGTGGTTTTTTG 1254

Db 4352 AAGAGTTGGTAGCTCTGTATCGCGCAAAACAAACACCGCTGTAGCGGTGGTTTTTTG 4411

QY 1255 TTTGCAAGCAGCAGATTACGCGCAGAAAAAAGGATCT 1292

Db 4412 TTTGCAAGCAGCAGATTACGCGCAGAAAAAAGGATCT 4449

RESULT 9

AR122288 AR122288 5865 bp DNA PAT 16-MAY-2001

LOCUS Sequence 8 from patent US 6165715.

DEFINITION AR122288

ACCESSION AR122288

VERSION AR122288.1 GI:14106605

KEYWORDS .

SOURCE Unknown.

ORGANISM Unknown.

REFERENCE 1 (bases 1 to 5865)

AUTHORS Collins,M.KatherineLevinge, Weiss,R.Anthony, Takeuchi,Y. and Cosset,F.

TITLE Expression systems

JOURNAL Patent: US 6165715-A 8 26-DEC-2000;

FEATURES

LOCATION/Qualifiers

1..5865

/organism="unknown"

BASE COUNT 1492 a 1602 c 1390 g 1373 t 8 others

ORIGIN

Query Match 33.4%; Score 604.4; DB 6; Length 5865;

Best Local Similarity 91.4%; Pred. No. 1.6e-164;

Matches 638; Conservative 0; Mismatches 60; Indels 0; Gaps 0;

QY 595 TTACTGGCTGTCTCCCTTATCGGGAAGCGGGCGGCATCATATCAAAATGACGCGCGCTG 654

Db 3752 TTATCATGTCTGGATCCAGATCTGGGCCCATGGCGCGGATCGATNNNNACATGTGAG 3811

QY 655 TAAAGTGTACGTTGAGAAGAAATTCCTGCAGCCCGCGGTTGCTGGCGTTTTTCCATA 714

Db 3812 CAAAAGGCCAGCAAAAGGCCAGGAACCGTAAAGAGCCGCTTGTGGCGTTTTTCCATA 3871

QY 715 GCGTCGCCCGCCCTGACGAGCATCACAAAATCGACGCTCAAGTCAGAGGTGGCGAAACC 774

Db 3872 GCGTCGCCCGCCCTGACGAGCATCACAAAATCGACGCTCAAGTCAGAGGTGGCGAAACC 3931

QY 775 CGACAGGACTATAAAGATACCAAGCGGTTTTCCCGCTGGAAGCTCCCTCGCGCTCTCCTG 834

Db 3932 CGACAGGACTATAAAGATACCAAGCGGTTTTCCCGCTGGAAGCTCCCTCGCGCTCTCCTG 3991

QY 835 TTCCGACCTGCGCGTTACGGATACCTGTCGCCCTTTCTCCCTTCGGAAGCGTGGCGC 894

Db 3992 TTCCGACCTGCGCGTTACGGATACCTGTCGCCCTTTCTCCCTTCGGAAGCGTGGCGC 4051

QY 895 TTCTCAATGCTCACGCTGTAGTATCTCAGTTTCGCTGAGGTCGTTCCCTCCAAAGCTGG 954

Db 4052 TTCTCAATGCTCACGCTGTAGTATCTCAGTTTCGCTGAGGTCGTTCCCTCCAAAGCTGG 4111

QY 955 GCTGTGTGCAGCAACCCCGGTTACGCCGACCGCTGCGCTTTATCCGGTAACTATCGTC 1014

Db 4112 GCTGTGTGCAGCAACCCCGGTTACGCCGACCGCTGCGCTTTATCCGGTAACTATCGTC 4171

QY 1015 TTGAGTCCAAACCCGGTAAAGACAGGACTTATCGCCACTGGCAGCAGCAGCTGGTAAACAGGA 1074

|||||
Db 4172 TTGAGTCCAAACCCGTAAGACACGACTTATCGCCACTGGCAGCACCCTGGTAAACAGGA 4231
QY 1075 TTACAGAGCAGGATGATAGCGGTGCTACAGAGTCTTGAAGTGGTGGCTTAACACTACG 1134
Db 4232 TTAGCAGAGCAGGATGATAGCGGTGCTACAGAGTCTTGAAGTGGTGGCTTAACACTACG 4291
QY 1135 GCTACACTAGAGGACAGTATTTGATCTGCGCTCTGCTGAAGCCAGTTACCTTCGGAA 1194
Db 4292 GCTACACTAGAGGACAGTATTTGATCTGCGCTCTGCTGAAGCCAGTTACCTTCGGAA 4351
QY 1195 AAAGAGTTGGTACGCTTGTATCGGCAAAACAAACACCGCTGGTAGCGGTGGTTTTTTG 1254
Db 4352 AAAGAGTTGGTACGCTTGTATCGGCAAAACAAACACCGCTGGTAGCGGTGGTTTTTTG 4411
QY 1255 TTGCAAGCAGCAGATTACGCGCCAGAAAAAAGGATCT 1292
Db 4412 TTGCAAGCAGCAGATTACGCGCCAGAAAAAAGGATCT 4449
RESULT 10
A60209 A60209 6028 bp DNA circular PAT 06-MAR-1998
LOCUS Sequence 5 from Patent WO9708330.
DEFINITION A60209
ACCESSION A60209.1 GI:3715217
VERSION
KEYWORDS
SOURCE
ORGANISM
unidentified.
unclassified.
REFERENCE 1 (bases 1 to 6028)
AUTHORS Collins, M.K., Weiss, R.A., Takeuchi, Y. and Cosset, F.
TITLE EXPRESSION SYSTEMS
JOURNAL Patent: WO 9708330-A 5 06-MAR-1997;
CANCER RES CAMPAIGN TECH (GB)
FEATURES
source
1. .6028
/organism="unidentified"
/db_xref="taxon:32644"
BASE COUNT 1515 a 1637 c 1499 g 1369 t 8 others
ORIGIN
Query Match 33.4%; Score 604.4; DB 6; Length 6028;
Best Local Similarity 91.4%; Pred. No. 1.6e-164;
Matches 638; Conservative 0; Mismatches 60; Indels 0; Gaps 0;
QY 595 TTACTGGCTGCTCCCTTATCGGGAAGCGGGCGCATCATATCAAAATGACGCGCGCTG 654
Db 3915 TTATCATGTCTGGATCCAGATCTGGGCCCATCGGCGCGGATCGATNNNNACATGTGAG 3974
QY 655 TAAAGTGTACGTTGAGAAGAATTCCTGCAGCCCGCGGTGCTGGCGTTTTCCTATA 714
Db 3975 CAAAAGGCCAGCAAAAGCCAGGAACCGTAAAGCCCGGTGCTGGCGTTTTCCTATA 4034
QY 715 GGTCTCGCCCTGACGAGCATCACAATAATCGAGCTCAAGTCAGAGTGGCGGAAC 774
Db 4035 GGCTCGCCCTGACGAGCATCACAATAATCGAGCTCAAGTCAGAGTGGCGGAAC 4094
QY 775 CGAGAGCTATAAAGATACAGCGGTTTCCCTCGGAAGTCCCTCGCGCTTCCTG 834
Db 4095 CGAGAGCTATAAAGATACAGCGGTTTCCCTCGGAAGTCCCTCGCTCCTCTG 4154
QY 835 TTCGACCTGCGCTTACCGGATACCTGTCGCCCTTCTCCCTTCGGGAAGCGTGGCG 894
Db 4155 TTCGACCTGCGCTTACCGGATACCTGTCGCCCTTCTCCCTTCGGGAAGCGTGGCG 4214
QY 895 TTTCTCAATGCTCAGCTGTAGGTATCTCAGTTCGGTGTAGGTTCGCTCCAAGCTG 954
Db 4215 TTTCTCAATGCTCAGCTGTAGGTATCTCAGTTCGGTGTAGGTTCGCTCCAAGCTG 4274
QY 955 GCTGTGTGACGAACCCCGCTTACGCCGACCGCTGCGCTTATCCGGTAACTATCGTC 1014
|||||

Db 4275 GCTGTGTGACGAACCCCGCTTTCAGCCGACCGCTCGGCTTATCCGGTAACTATCGTC 4334
QY 1015 TTGAGTCCAAACCCGTAAGACACGACTTATCGCCACTGGCAGCACCCTGGTAAACAGGA 1074
Db 4335 TTGAGTCCAAACCCGTAAGACACGACTTATCGCCACTGGCAGCACCCTGGTAAACAGGA 4394
QY 1075 TTACAGAGCAGGATGATAGCGGTGCTACAGAGTCTTGAAGTGGTGGCTTAACACTACG 1134
Db 4395 TTACAGAGCAGGATGATAGCGGTGCTACAGAGTCTTGAAGTGGTGGCTTAACACTACG 4454
QY 1135 GCTACACTAGAGGACAGTATTTGATCTGCGCTCTGCTGAAGCCAGTTACCTTCGGAA 1194
Db 4455 GCTACACTAGAGGACAGTATTTGATCTGCGCTCTGCTGAAGCCAGTTACCTTCGGAA 4514
QY 1195 AAAGAGTTGGTACGCTTGTATCGGCAAAACAAACACCGCTGGTAGCGGTGGTTTTTTG 1254
Db 4515 AAAGAGTTGGTACGCTTGTATCGGCAAAACAAACACCGCTGGTAGCGGTGGTTTTTTG 4574
QY 1255 TTGCAAGCAGCAGATTACGCGCCAGAAAAAAGGATCT 1292
Db 4575 TTGCAAGCAGCAGATTACGCGCCAGAAAAAAGGATCT 4612
RESULT 11
AR122285 AR122285 6028 bp DNA PAT 16-MAY-2001
LOCUS Sequence 5 from patent US 6165715.
DEFINITION AR122285
ACCESSION AR122285
VERSION AR122285.1 GI:14106602
KEYWORDS
SOURCE
ORGANISM
Unknown.
Unclassified.
REFERENCE 1 (bases 1 to 6028)
AUTHORS Collins, M.KatherineLevinge, Weiss, R.Anthony, Takeuchi, Y. and
Cosset, F.
TITLE Expression systems
JOURNAL Patent: US 6165715-A 5 26-DEC-2000;
FEATURES
source
1. .6028
/organism="unknown"
BASE COUNT 1515 a 1637 c 1499 g 1369 t 8 others
ORIGIN
Query Match 33.4%; Score 604.4; DB 6; Length 6028;
Best Local Similarity 91.4%; Pred. No. 1.6e-164;
Matches 638; Conservative 0; Mismatches 60; Indels 0; Gaps 0;
QY 595 TTACTGGCTGCTCCCTTATCGGGAAGCGGGCGCATCATATCAAAATGACGCGCGCTG 654
Db 3915 TTATCATGTCTGGATCCAGATCTGGGCCCATCGGCGCGGATCGATNNNNACATGTGAG 3974
QY 655 TAAAGTGTACGTTGAGAAGAATTCCTGCAGCCCGCGGTGCTGGCGTTTTCCTATA 714
Db 3975 CAAAAGGCCAGCAAAAGCCAGGAACCGTAAAGCCCGGTGCTGGCGTTTTCCTATA 4034
QY 715 GGTCTCGCCCTGACGAGCATCACAATAATCGAGCTCAAGTCAGAGTGGCGGAAC 774
Db 4035 GGCTCGCCCTGACGAGCATCACAATAATCGAGCTCAAGTCAGAGTGGCGGAAC 4094
QY 775 CGAGAGCTATAAAGATACAGCGGTTTCCCTCGGAAGTCCCTCGCGCTTCCTG 834
Db 4095 CGAGAGCTATAAAGATACAGCGGTTTCCCTCGGAAGTCCCTCGCTCCTCTG 4154
QY 835 TTCGACCTGCGCTTACCGGATACCTGTCGCCCTTCTCCCTTCGGGAAGCGTGGCG 894
Db 4155 TTCGACCTGCGCTTACCGGATACCTGTCGCCCTTCTCCCTTCGGGAAGCGTGGCG 4214
QY 895 TTTCTCAATGCTCAGCTGTAGGTATCTCAGTTCGGTGTAGGTTCGCTCCAAGCTG 954
Db 4215 TTTCTCAATGCTCAGCTGTAGGTATCTCAGTTCGGTGTAGGTTCGCTCCAAGCTG 4274

```
QY 955 GCTGTGTGCAGAACCCCGCTTTCAGCCGACCGCTGCGCTTATCCGGTAACTATGTC 1014
Db 4275 GCTGTGTGCAGAACCCCGCTTTCAGCCGACCGCTGCGCTTATCCGGTAACTATGTC 4334
QY 1015 TTGAGTCCAAACCCGGTAAGACACACTTATCGCCACTGGCAGCAGCCACTGGTAACAGGA 1074
Db 4335 TTGAGTCCAAACCCGGTAAGACACACTTATCGCCACTGGCAGCAGCCACTGGTAACAGGA 4394
QY 1075 TTAGCAGAGCAGGTATGTAGCGGTGCTACAGATTCTTGAAGTGTGCGCTTAACACTACG 1134
Db 4395 TTAGCAGAGCAGGTATGTAGCGGTGCTACAGATTCTTGAAGTGTGCGCTTAACACTACG 4454
QY 1135 GCTACACTAGAAGCAGAGTATTTGGTATCTGCGCTCTGCTGAAGCCAGTACCTTCGGAA 1194
Db 4455 GCTACACTAGAAGCAGAGTATTTGGTATCTGCGCTCTGCTGAAGCCAGTACCTTCGGAA 4514
QY 1195 AAGAGTTGGTAGCTCTTGATTCGGCGCAAAACAAACCCAGCGTGGTAGCGGTGTTTTTGG 1254
Db 4515 AAGAGTTGGTAGCTCTTGATTCGGCGCAAAACAAACCCAGCGTGGTAGCGGTGTTTTTGG 4574
QY 1255 TTTCGAAGCAGCAGATTACCGCGCAGAAAAAAGGATCT 1292
Db 4575 TTTCGAAGCAGCAGATTACCGCGCAGAAAAAAGGATCT 4612

RESULT 12
A60210
LOCUS A60210 6061 bp DNA circular PAT 06-MAR-1998
DEFINITION Sequence 6 from Patent WO9708330.
ACCESSION A60210
VERSION A60210.1 GI:3715218
KEYWORDS
SOURCE
ORGANISM
REFERENCE 1 (bases 1 to 6061)
AUTHORS Collins,M.K., Weiss,R.A., Takeuchi,Y. and Cosset,F.
TITLE EXPRESSION SYSTEMS
JOURNAL Patent: WO 9708330-A 6 06-MAR-1997;
CANCER RES CAMPAIGN TECH (GB)
FEATURES
source
location/Qualifiers
1..6061
/db_xref="taxon:32644"
/organism="unidentified"
BASE COUNT 1498 a 1665 c 1485 g 1405 t 8 others
ORIGIN

Query Match 33.4%; Score 604.4; DB 6; Length 6061;
Best Local Similarity 91.4%; Pred. No. 1.6e-164;
Matches 638; Conservative 0; Mismatches 60; Indels 0; Gaps 0;

QY 595 TTTACTGGCTGCTCCCTTATCGGGAAGCGGGCGCATCATATCAAAATGAGCGCGCTG 654
Db 3948 TTATCATGTCTGGATCCAGATCTGGGCCATCGGCCCGGATCGATNNNNACATGTGAG 4007
QY 655 TAAAGTGTAGTTGAGAAAGAAATTCGTGAGCGCGCGCGCTTGTGCTGCTTTTCCATA 714
Db 4008 CAAAGGCCAGCAAAAGCCAGAACCGTAAAGAGCGCGCTTGTGCTGCTTTTCCATA 4067
QY 715 GGCTCCGCCCGCTGACGAGCATCACAAAAATCGACGCTCAAGTCAGAGTGGCGAAACC 774
Db 4068 GGCTCCGCCCGCTGACGAGCATCACAAAAATCGACGCTCAAGTCAGAGTGGCGAAACC 4127
QY 775 GCACAGACTATAAAGATACAGCGCTTCCCGCTGGAAGCTCCCTCGTGGCTCTCCTG 834
Db 4128 GCACAGACTATAAAGATACAGCGCTTCCCGCTGGAAGCTCCCTCGTGGCTCTCCTG 4187
QY 835 TTCGACCCCTGCGCTTACCGGATACCTGTCCGCTTTCCTCCCTTCGGGAAGCGTGGCG 894
Db 4188 TTCGACCCCTGCGCTTACCGGATACCTGTCCGCTTTCCTCCCTTCGGGAAGCGTGGCG 4247
QY 895 TTTCTCAATGCTACGCTGTAGGTATCTCAGTTGCGGTGTAGGTGCTTCCGCTCCCAAGCTGG 954
```

```
Db 4248 TTTCTCAATGCTACGCTGTAGGTATCTCAGTTGCGGTGTAGGTGCTTCCGCTCCAGCTGG 4307
QY 955 GCTGTGTGCAGAACCCCGCTTTCAGCCGACCGCTGCGCTTATCCGGTAACTATGTC 1014
Db 4308 GCTGTGTGCAGAACCCCGCTTTCAGCCGACCGCTGCGCTTATCCGGTAACTATGTC 4367
QY 1015 TTGAGTCCAAACCCGGTAAGACACACTTATCGCCACTGGCAGCAGCCACTGGTAACAGGA 1074
Db 4368 TTGAGTCCAAACCCGGTAAGACACACTTATCGCCACTGGCAGCAGCCACTGGTAACAGGA 4427
QY 1075 TTAGCAGAGCAGGTATGTAGCGGTGCTACAGATTCTTGAAGTGTGCGCTTAACACTACG 1134
Db 4428 TTAGCAGAGCAGGTATGTAGCGGTGCTACAGATTCTTGAAGTGTGCGCTTAACACTACG 4487
QY 1135 GCTACACTAGAAGCAGAGTATTTGGTATCTGCGCTCTGCTGAAGCCAGTACCTTCGGAA 1194
Db 4488 GCTACACTAGAAGCAGAGTATTTGGTATCTGCGCTCTGCTGAAGCCAGTACCTTCGGAA 4547
QY 1195 AAGAGTTGGTAGCTCTTGATTCGGCGCAAAACAAACCCAGCGTGGTAGCGGTGTTTTTGG 1254
Db 4548 AAGAGTTGGTAGCTCTTGATTCGGCGCAAAACAAACCCAGCGTGGTAGCGGTGTTTTTGG 4607
QY 1255 TTTCGAAGCAGCAGATTACCGCGCAGAAAAAAGGATCT 1292
Db 4608 TTTCGAAGCAGCAGATTACCGCGCAGAAAAAAGGATCT 4645

RESULT 13
AR122286
LOCUS AR122286 6061 bp DNA PAT 16-MAY-2001
DEFINITION Sequence 6 from patent US 6165715.
ACCESSION AR122286
VERSION AR122286.1 GI:14106603
KEYWORDS
SOURCE
ORGANISM
REFERENCE 1 (bases 1 to 6061)
AUTHORS Collins,M.KatherineLevinge, Weiss,R.Anthony, Takeuchi,Y. and Cosset,F.
TITLE Expression systems
JOURNAL Patent: US 6165715-A 6 26-DEC-2000;
FEATURES
source
location/Qualifiers
1..6061
/organism="unknown"
BASE COUNT 1498 a 1665 c 1485 g 1405 t 8 others
ORIGIN

Query Match 33.4%; Score 604.4; DB 6; Length 6061;
Best Local Similarity 91.4%; Pred. No. 1.6e-164;
Matches 638; Conservative 0; Mismatches 60; Indels 0; Gaps 0;

QY 595 TTTACTGGCTGCTCCCTTATCGGGAAGCGGGCGCATCATATCAAAATGAGCGCGCTG 654
Db 3948 TTATCATGTCTGGATCCAGATCTGGGCCATCGGCCCGGATCGATNNNNACATGTGAG 4007
QY 655 TAAAGTGTAGTTGAGAAAGAAATTCGTGAGCGCGCGCGCTTGTGCTGCTTTTCCATA 714
Db 4008 CAAAGGCCAGCAAAAGCCAGAACCGTAAAGAGCGCGCTTGTGCTGCTTTTCCATA 4067
QY 715 GGCTCCGCCCGCTGACGAGCATCACAAAAATCGACGCTCAAGTCAGAGTGGCGAAACC 774
Db 4068 GGCTCCGCCCGCTGACGAGCATCACAAAAATCGACGCTCAAGTCAGAGTGGCGAAACC 4127
QY 775 GCACAGACTATAAAGATACAGCGCTTCCCGCTGGAAGCTCCCTCGTGGCTCTCCTG 834
Db 4128 GCACAGACTATAAAGATACAGCGCTTCCCGCTGGAAGCTCCCTCGTGGCTCTCCTG 4187
QY 835 TTCGACCCCTGCGCTTACCGGATACCTGTCCGCTTTCCTCCCTTCGGGAAGCGTGGCG 894
Db 4188 TTCGACCCCTGCGCTTACCGGATACCTGTCCGCTTTCCTCCCTTCGGGAAGCGTGGCG 4247
```

Qy	835	TTCCGACCGCTGGCGCTTACCGATACCTGTCCGCTTCTCCCTTCGGGAAGCGTGGCGC	894
Db	4439	TTCCGACCCCTGGCGCTTACCGGATACCTGTCCGCTTCTCCCTTCGGGAAGCGTGGCGC	4498
Qy	895	TTTCTCAATGCTCACGCTTAGGTATCTCAGTTGCGGTAGGTCGTCTCAAGCTGG	954
Db	4499	TTTCTCAATGCTCACGCTTAGGTATCTCAGTTGCGGTAGGTCGTCTCAAGCTGG	4558
Qy	955	GCTGTGTGCACGAACCCCGCTTACGCCGGACCGCTGGCCCTTATCCGGTAACCTATCGTC	1014
Db	4559	GCTGTGTGCACGAACCCCGCTTACGCCGGACCGCTGGCCCTTATCCGGTAACCTATCGTC	4618
Qy	1015	TTTGAGTCCAAACCGGTAAAGACAGCACTTATGCCACTGGCAGCAGCCACTGGTAACAGGA	1074
Db	4619	TTTGAGTCCAAACCGGTAAAGACAGCACTTATGCCACTGGCAGCAGCCACTGGTAACAGGA	4678
Qy	1075	TTAGCAGAGCGAGGTATGTAGCGGTGTCTACAGAGTCTTGTGAAGTGTGGCTTAACCTACG	1134
Db	4679	TTAGCAGAGCGAGGTATGTAGCGGTGTCTACAGAGTCTTGTGAAGTGTGGCTTAACCTACG	4738
Qy	1135	GCTACACTAGAAGCAGTATTTGGTATCTGGCGTCTGCTGAAGCCAGTTACCTTCGGAA	1194
Db	4739	GCTACACTAGAAGCAGTATTTGGTATCTGGCGTCTGCTGAAGCCAGTTACCTTCGGAA	4798
Qy	1195	AAACAGTTTGTAGTCTTTGATCCGGCAAAACAAACCAACCGCTGGTAGCGGTGGTTTTTG	1254
Db	4799	AAACAGTTTGTAGTCTTTGATCCGGCAAAACAAACCAACCGCTGGTAGCGGTGGTTTTTG	4858
Qy	1255	TTTCCAGCAGCAGATTTACCGCGAGAAAAAAGGATCT	1292
Db	4859	TTTCCAGCAGCAGATTTACCGCGAGAAAAAAGGATCT	4896

RESULT	15
AR122287	AR122287 6312 bp DNA PAT 16-MAY-2001
LOCUS	Sequence 7 from patent US 6165715.
DEFINITION	AR122287
ACCESSION	AR122287
VERSION	AR122287.1 GI:14106604
KEYWORDS	
SOURCE	Unknown.
ORGANISM	Unclassified.
REFERENCE	1 (bases 1 to 6312)
AUTHORS	Collins,M,KatherineLevinge, Weiss,R,Anthony, Takeuchi,Y., and Cosset,F.
TITLE	Expression systems
JOURNAL	Patent: US 6165715-A 7 26-DEC-2000;
FEATURES	Location/Qualifiers 1..6312 /organism="unknown"
BASE COUNT	1575 a 1692 c 1561 g 1476 t 8 others
ORIGIN	
Query Match	33.4%; Score 604.4; DB 6; Length 6312;
Best Local Similarity	91.4%; Pred.No.1.6e-104; Indels 0; Gaps:
Matches	638; Conservative
QY	595 TTTACTGGCTGCTCCCTTATCGGAAGCGGGCGCATCATCAAAATGACGCGCGCTG 659
Db	4199 TTATCATGTCTGGATCCAGATCTGGCCCCATCGCGCGCGATCGATNNNACATGTGAG 429
QY	655 TAAAGTGTTACGTGTGAAGAAGATTCTGCAGCACC CGCGCGCTTTCCTGGCTTTTCCATA 715
Db	4259 CAAAAGGCCAGCAAAGGGCGAGAACCGTTAAAGAGGGCGCGTCTGCTGGCGTTTTTCCATA 4359
QY	715 GGCTTCGCGCCCCTGACGAGCATCACAAAAATCGACGCTCAAGTCAGAGGTGCGCAAACC 775
Db	4319 GGCTTCGCGCCCCTGACGAGCATCACAAAAATCGACGCTCAAGTCAGAGGTGCGCAAACC 4379
QY	775 GCACGAGGACTATAAAGATACCGAGGGGTTTTCCCCCTGGAGCTCCCTGTGGGCTCTCCTG 839

Db 4379 CGACAGACTATATAAGATACACGCGGTTTCCCGCTGGAAGTCCCTCGTGCCTCTCCTG 4438
QY 835 TTCCGACCCCTGCGCTTACCGGATACCTGTCCGCTTTCTCCCTTCGGGAAGCGTGGCGC 894
Db 4439 TTCCGACCCCTGCGCTTACCGGATACCTGTCCGCTTTCTCCCTTCGGGAAGCGTGGCGC 4498
QY 895 TTTCTCAATGCTCACGCTGTAGGTATCTCAGTTTCGGTGTAGGTTCGGTTCGCTCCAAAGCTGG 954
Db 4499 TTTCTCAATGCTCACGCTGTAGGTATCTCAGTTTCGGTGTAGGTTCGGTTCGCTCCAAAGCTGG 4558
QY 955 GCTGTGTGACAGAACCCCGTTTCAGCCCGACCGCTGCGGCTTATCCGGTAACTATCGTC 1014
Db 4559 GCTGTGTGACAGAACCCCGTTTCAGCCCGACCGCTGCGGCTTATCCGGTAACTATCGTC 4618
QY 1015 TTGAGTCCAAACCCGGTAAGACACGACTTATCGGCACCTGGCAGCAGCCACTGGTAAACAGGA 1074
Db 4619 TTGAGTCCAAACCCGGTAAGACACGACTTATCGGCACCTGGCAGCAGCCACTGGTAAACAGGA 4678
QY 1075 TTAGCAGAGCGAGGTATGTAGCGGGTGTACAGAGTTCTTGAAGTGGTGGCCCTAACTACG 1134
Db 4679 TTAGCAGAGCGAGGTATGTAGCGGGTGTACAGAGTTCTTGAAGTGGTGGCCCTAACTACG 4738
QY 1135 GCTACACTAGAAGCACAGTATTTGGTATCTGCGCTGCTGCTGAAGCCAGTTACCTTCGGAA 1194
Db 4739 GCTACACTAGAAGCACAGTATTTGGTATCTGCGCTGCTGCTGAAGCCAGTTACCTTCGGAA 4798
QY 1195 AAAGAGTTGGTAGCTCTTGATCCGGCAAAACACACCGCTGGTGGTGGTGGTTTTTTTG 1254
Db 4799 AAAGAGTTGGTAGCTCTTGATCCGGCAAAACACACCGCTGGTGGTGGTGGTTTTTTTG 4858
QY 1255 TTGCAAGCAGCAGATTACCGCAGAAAAAAGGATCT 1292
Db 4859 TTGCAAGCAGCAGATTACCGCAGAAAAAAGGATCT 4896

THIS PAGE BLANK (USPTO)

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: January 17, 2002, 12:04:29 ; Search time 599.86 Seconds
(without alignments)
2582.580 Million cell updates/sec

Title: US-09-242-202A-28

Perfect score: 1807
Sequence: 1 GGTACCTGCCACCATGGCGC.....CTCCACGAGCTCTCCACA 1807

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 930621 seqs, 428662619 residues

Total number of hits satisfying chosen parameters: 1861242

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : N_Geneseq_1101.*
1: /SIDS2/gcgdata/geneseq/geneseq/NA1980.DAT.*
2: /SIDS2/gcgdata/geneseq/geneseq/NA1981.DAT.*
3: /SIDS2/gcgdata/geneseq/geneseq/NA1982.DAT.*
4: /SIDS2/gcgdata/geneseq/geneseq/NA1983.DAT.*
5: /SIDS2/gcgdata/geneseq/geneseq/NA1984.DAT.*
6: /SIDS2/gcgdata/geneseq/geneseq/NA1985.DAT.*
7: /SIDS2/gcgdata/geneseq/geneseq/NA1986.DAT.*
8: /SIDS2/gcgdata/geneseq/geneseq/NA1987.DAT.*
9: /SIDS2/gcgdata/geneseq/geneseq/NA1988.DAT.*
10: /SIDS2/gcgdata/geneseq/geneseq/NA1989.DAT.*
11: /SIDS2/gcgdata/geneseq/geneseq/NA1990.DAT.*
12: /SIDS2/gcgdata/geneseq/geneseq/NA1991.DAT.*
13: /SIDS2/gcgdata/geneseq/geneseq/NA1992.DAT.*
14: /SIDS2/gcgdata/geneseq/geneseq/NA1993.DAT.*
15: /SIDS2/gcgdata/geneseq/geneseq/NA1994.DAT.*
16: /SIDS2/gcgdata/geneseq/geneseq/NA1995.DAT.*
17: /SIDS2/gcgdata/geneseq/geneseq/NA1996.DAT.*
18: /SIDS2/gcgdata/geneseq/geneseq/NA1997.DAT.*
19: /SIDS2/gcgdata/geneseq/geneseq/NA1998.DAT.*
20: /SIDS2/gcgdata/geneseq/geneseq/NA1999.DAT.*
21: /SIDS2/gcgdata/geneseq/geneseq/NA2000.DAT.*
22: /SIDS2/gcgdata/geneseq/geneseq/NA2001.DAT.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description	
1	1807	100.0	1807	19	AAV21733 Humanised vector p	
2	1583.2	87.6	2308	19	AAV21734 Humanised vector p	
3	1310.4	72.5	1547	19	AAV21732 Humanised vector p	
c	4	755	41.8	2932	13	AAQ21163 COS cell expressio
c	5	755	41.8	2932	17	AAQ38557 Expression vector
c	6	755	41.8	2932	17	AAQ14702 Vector plasmid pH
c	7	755	41.8	2932	19	AAV63441 Expression vector
c	8	755	41.8	2932	20	AAV81199 Expression vector
c	9	755	41.8	2932	21	AAV50577 Eukaryotic cell ex
c	10	755	41.8	2932	22	AAV503171 Expression vector
11	607.4	33.6	7252	20	AAV86928 Complete sequence	

12	604.4	33.4	5865	18	AAV90692	Plasmid FdelPRDSA
13	604.4	33.4	6028	18	AAV90689	Plasmid FdelPMOAF
14	604.4	33.4	6061	18	AAV90690	Plasmid FdelPMOSA
15	604.4	33.4	6312	18	AAV90691	Plasmid FdelPMGASA
16	604.2	33.4	7073	18	AAV50962	TF8-5G9 CDR-grafte
17	604.2	33.4	7864	18	AAV50963	TF8-5G9 CDR-grafte
18	604.2	33.4	11795	21	AAV59345	Nucleotide sequenc
19	604.2	33.4	13254	17	AAV40915	Nucleotide sequenc
20	604.2	33.4	13254	21	AAV32151	pEel2 Combo BM 12
21	604.2	33.4	13254	21	AAV32151	Complete nucleotid
22	603	33.4	608	19	AAV21731	ColE1 origin of re
23	603	33.4	2077	20	AAV23771	Vector pASlib DNA.
24	603	33.4	2462	21	AAV74638	Plasmid pSP72. Es
25	603	33.4	2577	14	AAQ36620	Expression plasmid
26	603	33.4	2704	20	AAQ206758	Vector pUC28 nucle
27	603	33.4	2927	17	AAV29158	Plasmid pTRP. Syn
28	603	33.4	2927	17	AAV31789	Plasmid pTRP. Syn
29	603	33.4	2939	16	AAQ87350	Plasmid BGINV. Sy
30	603	33.4	2939	19	AAV37293	pBGINV plasmid use
31	603	33.4	3003	11	AAQ05745	Plasmid P_L-mu-smc
32	603	33.4	3104	10	AAV90296	DNA target sequenc
33	603	33.4	3130	19	AAV46334	Universal reporter
34	603	33.4	3175	20	AAV220066	DNA encoding gluta
35	603	33.4	3327	22	AAV86263	Plasmid GHRH1-29YA
36	603	33.4	3331	20	AAV84434	pBC56-01, containi
37	603	33.4	3331	21	AAV299245	Nucleotide sequenc
38	603	33.4	3344	21	AAV39495	Transgenic unc-119
39	603	33.4	3369	22	AAV86255	pGHRH-4 44SK const
40	603	33.4	3403	9	AAV80956	Plasmid pDS5/RBSII
41	603	33.4	3414	12	AAQ12785	pDS56/RBSII-2 sequ
42	603	33.4	3414	14	AAQ34609	Expression plasmid
43	603	33.4	3415	12	AAQ12784	pDS56/RBSII-1 sequ
44	603	33.4	3415	14	AAQ34608	Expression plasmid
45	603	33.4	3416	9	AAV80958	Plasmid pDS56/RBSI

ALIGNMENTS

RESULT 1	AAV21733	standard; cDNA; 1807 BP.
ID	AAV21733	
XX	AAV21733	
XX	AAV21733	
DT	17-AUG-1998	(first entry)
XX	Humanised vector	pITL-1.
DE	Vector;	vaccine; tumour; antigen; plasmid pITL-1; ds.
XX	Chimeric - Homo sapiens.	
OS	Chimeric - Escherichia coli.	
OS	Synthetic.	
XX	Key	Location/Qualifiers
XX	CDS	472..680
FT		/*tag= C
FT		/note= "SupF gene"
FT	misc_feature	686..1292
FT		/*tag= d
FT		/note= "ColE1 origin of replication"
FT	promoter	1605..1847
FT		/*tag= e
FT		/note= "RANTES promoter"
XX	WO9806863-A1.	
XX	19-FEB-1998.	
XX	14-AUG-1997;	97WO-US14306.
XX	14-AUG-1996;	96US-0023931.

(USSH) US DEPT HEALTH & HUMAN SERVICES.

Nelson EL, Nelson PJ;

WPI; 1998-159552/14.

Humanised polynucleotide vectors - comprising human derived promoter and sequence acceptance site, used for the production of vaccines

and sequence acceptance site, used for the production of vaccines

Claim 14; Page 56; 125pp; English.

Plasmid pITL-1 comprises a base vector for novel humanised

polynucleotide vectors. The vector in *Escherichia coli* DH

10-beta/P3 is deposited as ATCC 98400. Novel vectors comprise a

human-derived promoter or mammalian homologue which is functional

human derived promoter or mammalian homologue which is functional in mammalian target tissue and cells and a sequence acceptance site

in mammalian target tissue and cells and a sequence acceptance site (see AAV21735-36) which accepts cDNA products from RT-PCR cloning.

They also contain minimal non-human components, such as a replication (see 11/21/95 30) which accepts cDNA products from RT-PCR cloning.

they also contain minimal non-human components, such as a replication origin (see AAV21715) and selectable marker gene (see AAV21717-18) that bat

cityin (see AAV21/13) and selectable marker gene (see AAV21/1/-18) that are necessary for production of the vector. The novel vectors are

are necessary for production of the vector. The novel vectors are used to express target antigens, especially tumour antigens, thereby

used to express target antigens, especially tumour antigens. They are non-replicating in mammalian cells but are capable of extended

are non-replicating in mammalian cells but are capable of extended stable expression of target sequences generating an immune response

stable expression of target sequences generating an immune response in immunised individuals. The vectors selectively elicit immune

in immunised individuals. The vectors selectively elicit immune responses to the target sequences with little or no immune response

Responses to the target sequences with little or no immune response to the other components of the vectors. The target antigens are

to the other components of the vectors. The target antigens are expressed as intracellular polypeptides or peptides and as such

expressed as intracellular polypeptides or peptides and, as such, are processed as self polypeptides or peptides and secreted into

are processed as self polypeptides or peptides and appropriately presented on antigen presenting cells

Sequence 1807 BP: 410 A: 463 C: 403 C: 423 E: 0 other:

Query Match

Query Match 100.0%; Score 1807; DB 19; Length 1807;

Best Local Similarity 100.0%; Pred, No. 0;

1 GGTACCTGCCACCACCATGGCGCGGATTCCTTATCACTGATAAGTTGGTGGACATATATATGTT 60

1 ggtacctgccaccatggcgcggattctttatcactgaataagtttgtgacatatatgtt 60

51 TATCAGTGATAAAGTGTCAAGCATGACAAAGTTGCAGCGCGAATACAGTGATCCGTGCGCG 120

61 tatcagtgataaagtgtcaaacatcaccaagttagcaaacgaatcacctacatccatgacggg

121 CCTGGACTGTTGAACCGCTCCGCCGTGACCCCTCTCACCACCCCAAGCC

[illegible][illegible]

101

[illegible]

301

100

361

362

363

364

365

366

367

368

369

370

371

372

373

374

375

376

377

378

379

380

381

382

383

384

385

386

387

388

389

390

391

392

393

394

395

396

397

398

399

400

401

402

403

404

405

406

407

408

409

410

411

412

413

414

415

416

417

418

419

420

421

422

423

424

425

426

427

428

429

430

431

432

433

434

435

436

437

438

439

440

441

442

443

444

445

446

447

448

449

450

451

452

453

454

455

456

457

458

459

460

461

462

463

464

465

466

467

468

469

470

471

472

473

474

475

476

477

478

479

480

481

482

483

484

485

486

487

488

489

490

491

492

493

494

495

496

497

498

499

500

501

502

503

504

505

506

507

508

509

510

511

512

513

514

515

516

517

518

519

520

521

522

523

524

525

526

527

528

529

530

531

532

533

534

535

536

537

538

539

540

541

542

543

544

545

546

547

548

549

550

551

552

553

554

555

556

557

558

559

560

561

562

563

564

565

566

567

568

569

570

571

572

573

574

575

576

577

578

579

580

581

582

583

584

585

586

587

588

589

590

591

592

593

594

595

596

597

598

599

600

601

602

603

604

605

606

607

608

609

610

611

612

613

614

615

616

617

618

619

620

621

622

623

624

625

626

627

628

629

630

631

632

633

634

635

636

637

638

639

640

641

642

643

644

645

646

647

648

649

650

651

652

653

654

655

656

657

658

659

660

661

662

663

664

665

666

667

668

669

670

671

672

673

674

675

676

677

678

679

680

681

682

683

684

685

686

687

688

689

690

691

692

693

694

695

696

697

698

699

700

701

702

703

704

705

706

707

708

709

710

711

712

713

714

715

716

717

718

719

720

721

722

723

724

725

726

727

728

729

730

731

732

733

734

735

736

737

738

739

740

741

742

743

744

745

746

747

748

749

750

751

752

753

754

755

756

757

758

759

760

761

762

763

764

765

766

767

768

769

770

771

772

773

774

775

776

777

778

779

780

781

782

783

784

785

786

787

788

789

790

791

792

793

794

795

796

797

798

799

800

801

802

803

804

805

806

807

808

809

810

811

812

813

814

815

816

817

818

819

820

821

822

823

824

825

826

827

828

829

830

831

832

833

834

835

836

837

838

839

840

841

842

843

844

845

846

847

848

849

850

851

852

853

854

855

856

857

858

859

860

861

862

863

864

865

866

867

868

869

870

871

872

<

20

[illegible]

100

.....

[illegible]

```
QY 1621 GCATTTTCAGTTTCTTTTCCGTTTGTGCAATTCACATTATGATACCGGCCCAATGCTTG 1680
Db 1621 gctatttcagtttcttttccgtttgtgcaatttcacattatgataccggccaatgcttg 1680
QY 1681 GTTGCTATTTCGAAACTCCCTTAGGGGATGCCCTCAACTGGCCCTATAAGAGGCCAG 1740
Db 1681 gtgctatttggaaactcccttagggatgcccctcaactgcccataaaaggccag 1740
QY 1741 CTTGAGCTGCAGAGGATTCCTGCAGAGGATCAAGACAGCACGTTGGACCTCGCACAGCCTC 1800
Db 1741 cctgagctgcagagattcctgcagagatcaagacagcacgctggacctgcacagcctc 1800
QY 1801 TCCACA 1807
Db 1801 tccaca 1807

RESULT 2
AAV21734
ID AAV21734 standard; cDNA; 2308 BP.
AC AAV21734;
XX
XX 17-AUG-1998 (first entry)
DE Humanised vector pITL-1 GFP.
KW Vector; vaccine; tumour; antigen; plasmid pITL-1 GFP;
KW green fluorescent protein; ds.
XX
OS Chimeric - Homo sapiens.
OS Chimeric - Escherichia coli.
OS Chimeric - Aequorea victoria.
OS Synthetic.
XX
FH Key
FT 20..734
FT CDS
FT
FT /*tag= a
FT /product= green fluorescent protein
FT 973..1181
FT /*tag= c
FT /note= "SupF gene"
FT 1191..1793
FT /*tag= d
FT /note= "ColE1 origin of replication"
FT 2063..2308
FT /*tag= e
FT /note= "RANTES promoter"
XX
XX W09806863-A1.
XX
XX 19-FEB-1998.
XX
XX 14-AUG-1997; 97WO-US14306.
XX
XX 14-AUG-1996; 96US-0023931.
XX
XX (USSH ) US DEPT HEALTH & HUMAN SERVICES.
XX
XX Nelson EL, Nelson PJ;
XX
XX WPI; 1998-159552/14.
XX
XX Humanised polynucleotide vectors - comprising human derived promoter
XX and sequence acceptance site, used for the production of vaccines
XX
XX Example 11; Page 57-58; 125pp; English.
XX
XX Plasmid pITL-1 GFP comprises base vector pITL-1 (see AAV21733) and
XX a humanised green fluorescent protein (GFP) reporter sequence
XX (see AAV21725). Novel humanised vectors comprise a human-derived
XX promoter or mammalian homologue which is functional in mammalian
XX target tissue and cells and a sequence acceptance site which
```

```
CC accepts cDNA products from RT-PCR cloning. The novel vectors are
CC used to express target antigens, especially tumour antigens. They
CC are non-replicating in mammalian cells but are capable of extended
CC stable expression of target sequences generating an immune response
CC in immunised individuals. The vectors selectively elicit immune
CC responses to the target sequences with little or no immune response
CC to the other components of the vectors. The target antigens are
CC expressed as intracellular polypeptides or peptides and, as such,
CC are processed as self polypeptides or peptides and appropriately
CC presented on antigen presenting cells.
XX
XX Sequence 2308 BP; 571 A; 601 C; 604 G; 532 T; 0 other;
SQ

Query Match 87.6%; Score 1583.2; DB 19; Length 2308;
Best Local Similarity 99.8%; Pred. No. 0;
Matches 1585; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 220 GAACAAAGCGGGCGCCTTAAGGGCCATATGTGATGGATCCTTGACCCAGCGGGGAT 279
Db 721 gtacaagtgaagcgcccttaaggccatatgtgagtggatgacctgagccagcggggat 780
QY 280 GGGGGAGACCTGTAGTCAGAGACCCCGGGGAGACAGAGCCCAATGCCGCTCTTCCCTGC 339
Db 781 gggggagacctgtagtcagagaccccgggcagcagccaatgccgctcttccctgc 840
QY 340 AGGATGAGTAGTGAGTGCCCTCTCTGGCCCTGGAAGTGCACATCCACACACACGCC 399
Db 841 aggatgagtgtgagtgcctctcctggccctggaagtgcacactccagtgccaccagcc 900
QY 400 TTGCTCTAATAAAATTAAGTTGTCATCATTTTGTCTGACTAGTGCTCTATATATATAT 459
Db 901 ttgtcctaataaataaagtgtcatcatcttctgactaggtgtcctataatattat 960
QY 460 AAGCTTGATATCGAATCTTTTCGACTTTTGAAGTGATGTGTGGGGGAAGGATTCGA 519
Db 961 aagcttgatatacgaattcttctcgacttttgaagtgtggtggtgggaaggattcga 1020
QY 520 ACCTTCGAAGTCGATGACGGCAGATTTAGAGTCTGCTCCCTTTGGCGCGCTCGGGAACCCC 579
Db 1021 accttcgaagctgatgacggcagatttagagtctgctcccttggccgctcgggaaacccc 1080
QY 580 ACCACGGGTAAATGCTTTTACTGGCCTGCTCCCTTATCGGAAGCGGGCGCATCATATCA 639
Db 1081 accacgggtaagtcttttactggcctgtcctctatcggaagcgggcgcatcatca 1140
QY 640 AATGACGCGCGCTGTAAAGTGTACGTTGAGAAAGAAATTCCTGCACCCCGCGCGTTGC 699
Db 1141 aatgacgcgcgctgtaaagtgttacgttgagaaagaattcttgacgcccgcgctgtgc 1200
QY 700 TGGCGTTTTCATAGGCTCCGCCCTCGACGAGCATCAAAAATCGACGCTCAAGTC 759
Db 1201 tggcgttttccataggctccgccccctgacgagcatcaaaaaatcgacgctcaagtc 1260
QY 760 AGAGGTGGCGAAACCCGACAGGACTATAAGATACACAGGGGTTTCCCGCTGGAAAGCTCCC 819
Db 1261 agaggtggcgaaacccgacagactataagataccagcggtttccctcggaagctccc 1320
QY 820 TCGTGGCTCTCTGTTCGACCGCTGCGCGTACCGGATACCTGCGCGCTTTCTCCCTT 879
Db 1321 tcgtgctctctctgttccgacctgcgcttaccggatacctgtccgcttctctcctt 1380
QY 880 CGGGAAGCGTGGCGCTTTCTCAATGCTCAGCTAGTATCTCAGTTCCGTGTAGTGTGCG 939
Db 1381 cgggaagcgtggcgcttctcctcaatgctcagcgtgtaggtatctcagttcgtgtgagtcg 1440
QY 940 TTGCTCTCAAAGCTGGGCTGTGTGCACGAACCCCGCTTCAGCCCGACCGCTGCGCTTAT 999
Db 1441 ttgcgtccaagctgggctgtgtgcagaaacccccggttcagccgacgctgaccttat 1500
QY 1000 CCGGTAACTATCGTCTTGTAGTTCACCCGGTGAAGACACAGACTTATCGCTGCGCAG 1059
Db 1501 ccggttaactatgctcttgagttccaaaccccggttaagacacgacttatccgacctgcagcag 1560
```

FT	misc_feature	/note= "SupF gene"
FT	686..1292	
FT	/*tag= d	
FT	/note= "Cole1 origin of replication"	
FT	1311..1547	
FT	promoter	
FT	/*tag= e	
FT	/note= "RANTES promoter"	
XX	W09806863-AL.	
PN		
XX	19-FEB-1998.	
PD		
XX	14-AUG-1997;	97WO-US14306.
PF		
XX	14-AUG-1996;	96US-0023931.
PR		
XX	(USSH) US DEPT HEALTH & HUMAN SERVICES.	
XX	Nelson EL, Nelson PJ;	
PI		
XX	WPI; 1998-159552/14.	
DR		
XX	Humanised polynucleotide vectors - comprising human derived promoter and sequence acceptance site, used for the production of vaccines	
FT		
FT		
XX	Claim 14; Page 56-57; 125pp; English.	
PS		
XX	Plasmid pITL-A comprises a base vector for novel humanised polynucleotide vectors. The vector in Escherichia coli DH 10-beta/P3 is deposited as ATCC 98401. Novel vectors comprise a human-derived promoter or mammalian homologue which is functional in mammalian target tissue and cells and a sequence acceptance site (see AAV21735-36) which accepts cDNA products from RT-PCR cloning. They also contain minimal non-human components, such as a replication origin (see AAV21715) and selectable marker gene (see AAV21717-18) that are necessary for production of the vector. The novel vectors are used to express target antigens, especially tumour antigens. They are non-replicating in mammalian cells but are capable of extended stable expression of target sequences generating an immune response in immunised individuals. The vectors selectively elicit immune responses to the target sequences with little or no immune response to the other components of the vectors. The target antigens are expressed as intracellular polypeptides or peptides and, as such, are processed as self polypeptides or peptides and appropriately presented on antigen presenting cells.	
XX	Sequence 1547 BP; 344 A; 408 C; 425 G; 370 T; 0 other;	
SQ		
Query Match 72.5%; Score 1310.4; DB 19; Length 1547;		
Best Local Similarity 99.9%; Pred. No. 0;		
Matches 1311; Conservative 0; Mismatches 1; Indels 0; Gaps 0;		
QY	1	GGTACCTGCCACCATGGCGGGATTCTTTTATCTACTGATTAAGTTGGTGACATATTATGTT 60
Db	1	ggtaacctgccaccatggcgggattcttttatactgataagttggcgacatatattggtt 60
QY	61	TATCAGTGATAAAGTGTCAAGCATGACAAAGTTGCGAGCCGAATACAGTATCGTGCCCGG 120
Db	61	tatcagtgtataaagtgccaagcatgacaaagtgcagcgaatacacagtatcgctgccgg 120
QY	121	CCCTGGGACTGTGAACGAGGTGCGCGTAGACGGTCTGCACGACGCAAACTGCGGGAACG 180
Db	121	ccctgggactgtgaacgaggtgcgcgtagacggtctgcacgacgcaaaactgcggaacg 180
QY	181	GTTGGGGGTGCAGCAGCGCGGCTTTTACTGCACTTCAGGAACAAGCGCGCTTAAGG 240
Db	181	gttgggggtgcagcagcgcgcttttactggcacttcaggaacaagcgcgcttaagg 240
QY	241	GCCATATGGTGAGTGGATGCGCTTGACCCCGAGCGGGGATGGGGAGACCTGTAGTCAGAG 300
Db	241	gccatatggtgagtgagtgatgccttgaccccgagcggggatggggagacctgtagtcagag 300

QY 301 CCCCGGCGCAGCACAGGCAATGCCGCTCTCCCTGCAGGATGAGTAGTGCCCTC 360
DB 301 ccccgggcagcacagggccaaTgcgcgtccttccccctcagatgagtagtgccctc 360
QY 361 TCCCTGGCCCTGGAAGTTGCCACTCCAGTCCGCCACACAGCCTTGCTCTAATAAAATTAAGTT 420
DB 361 tccctggccctggaagtgcactccagtgccaccagcctgtccctaaTaaaaattaagtt 420
QY 421 GCATCATTTTGTCTGACTAGGTGCTCTCTATAATATTAAGCTTGATATCGAATTCCTTT 480
DB 421 gcatcatTTTgtctgactaggTgtctctataataTataagcttgatatacgaattcttt 480
QY 481 CGGACTTTTGAAGTGATGGTGGTGGGGAAGGATTCGAACCTTCGAAGTCGAGTACGGC 540
DB 481 cggactTTTgaagtgtatggtggTgggggaaggattcgaaaccttcgaagtgcgatgacggc 540
QY 541 AGATTTAGAGTGTGCTCTTTGGCGCGTCCGGGAACCCACACAGGGTAAATGCTTTTACT 600
DB 541 agatttagagTctgtcccttTggccgctcgggaaacccaccacgggtaagtgtttact 600
QY 601 GGCGTCTCCCTATTCGGAAGCGGGGCGCATCATATCAATGACGCGCGCTGTAAAGT 660
DB 601 ggcgtctccctatctcggaagcgggggcgcatcataTcaaatgacgcgcgctgtaaagt 660
QY 661 GTTACGTTGAGAAAGAAATTCCTGCAGCGCGCGTGTGCTGGCGTTTTCCTATAGGCTCC 720
DB 661 gttacgttgagaaagaattcctgcagccgcgcgcgtgtgctggcgtttttccatagagctcc 720
QY 721 GCGCCCTGACGAGCATACAAAATCGAGCTCAAGTCAGAGGTGGCGAAACCCGACAG 780
DB 721 gccccctgacgagcatcacaaaaatcgagctcaagtcagagtgTgcgaaccccgacag 780
QY 781 GACTATAAGATACGAGCGGTTCCCTCCCTGGAAGCTCCCTCGTGCCTCTCTCTGTTCCGA 840
DB 781 gactataaagataccagggtttcccccctggaaagctccctcgtgcgtctcctgttccga 840
QY 841 CCGTCGCGTTACCGGATACCTGTCCGCGCTTCTCCCTTCGGGAAGCGTGGCGTTTCTC 900
DB 841 cctgcgcgttaccggataacctgtccgcgtttctcccttcgggaagcgtggcgcttctc 900
QY 901 AATGCTACGCTGTAGGTATCTAGTTTCGGTGTAGTGTGCTGCTCCAAAGCTGGGTG 960
DB 901 aatgctcacgctgtaggtatctcagttcgggtgtaggtcgttcgctccaaagctgggctgtg 960
QY 961 TGCACCAACCCCGCTACGCGACCGCTGCGCTTATCCGCTAACTATCGTCTTGAAGT 1020
DB 961 tgcaacgaaccccccgtctcagccgcgcgcgttcccttataccggttaactcgtcttgagt 1020
QY 1021 CCAACCCGGTACACACGACTTATCGCCACTGGCAGCAGCCACTGGTAACAGGATTAGCA 1080
DB 1021 ccaacccggtaaagacagacttatcgccactggcagcagccactgTtaacaggattagca 1080
QY 1081 GAGCGAGGTATGAGCGGTGTACAGAGTTCCTGAAGTGGTGGCTTAACACTACGGGTACA 1140
DB 1081 gagcgaggtatgtagcggtgtctacagagttcttgaagtggTggcctaactacgcgtaca 1140
QY 1141 CTAGAGGACAGATTTGGTATCTGCGCTGCTGCTGAAGCCAGTTACTTCGGAAAAAGAG 1200
DB 1141 ctagaaggacagtatTTggtatctgcgcctgctgagccagttactcttcggaaaaagag 1200
QY 1201 TTGGTAGCTCTTGATCCGGCAAAACACACCGCTGGTAGCGGTGGTGTGTTTGTGTTGCA 1260
DB 1201 ttggtagctcttgatccggcaaaacacacccgctggtgagcggTggtttttgtttgtgca 1260
QY 1261 AGCAGCAGATTACGCGCAGAAAAAGGATCTGGGGGATCCGGAGAGCTCCC 1312
DB 1261 agcagcagattacgcgcagaaaaaaggatctgggggatccggagagctcac 1312

RESULT 4

AAQ21163/c

ID AAQ21163 standard; cDNA; 2932 BP.

XX

AC AAQ21163;
XX 21-MAY-1992 (first entry)
XX COS cell expression vector pIH3.
DE cloning technique; CD2; cell surface antigen; Immunodiagnosis;
KW tumour; ss.
XX Synthetic.
OS
FH Key Location/Qualifiers
FT misc_feature 1..589
FT /tag= a
FT /function= ori
FT /note= "derived from pMB1 origin"
FT misc_feature 590..597
FT /tag= b
FT /function= linker
FT /note= "SacII linker"
FT misc_feature 598..799
FT /tag= c
FT /label= supF_gene
FT /note= "derived from synthetic tyrosine suppressor
trNA gene"
FT misc_feature 800..947
FT /tag= d
FT /label=
FT /note= "remnant of ASV LTR fragment (PvuII to MluI)"
FT misc_feature 948..1500
FT /tag= e
FT /function= enhancer
FT /note= "derived from cytomegalovirus Ad169 enhancer"
FT misc_feature 1501..1650
FT /tag= f
FT /note= "derived from HIV TATA and tat responsive-
elements"
FT misc_feature 1651..1716
FT /tag= g
FT /function= polylinker
FT /note= "derived from pILNAN polylinker"
FT misc_feature 2570..2917
FT /tag= h
FT /function= ori
FT /note= "derived from SV40 origin of replication"
FT misc_feature 2918..2922
FT /tag= i
FT /note= "derived from pIX, remnant of R1 site
from polylinker"
FT misc_signal 1717..2569
FT /tag= j
FT /note= "derived from pSV to splice and
poly-A addition signals"
FT
XX WO9201049-A.
XX 23-JAN-1992.
XX 15-JUL-1990; 90WO-US04986.
XX 13-JUL-1990; 90US-0553759.
XX (GEO-) GEN HOSPITAL CORP.
XX Seed B, Aruffo A, Amiot M;
XX WPI; 1992-056864/07.
XX New CD53 cell surface antigen and DNA encoding it - for
immuno-therapy and diagnosis of haematopoietic neoplasms, etc.
XX Disclosure; Fig 1; 160pp; English.
XX

Fri Jan 18 08:27:46 2002

This COS cell expression vector was constructed from pISV (Little et al., Mol. Biol. Med. 1:473-488 (1983)) by inserting a synthetic transcription unit between the suppressor tRNA gene and the SV40 origin. The transcription unit consisted of a chimeric promoter composed of human cytomegalovirus AD169 immediate early enhancer sequences fused to the HIV LTR -67 to (plus)80 sequences. A polylinker contg. 2 BstXI sites separated by a 350bp stuffer was inserted immediately downstream from the LTR (plus)80 sequence. Downstream from the polylinker were placed the SV40 small t antigen splice and early region polyadenylation signals derived from pSV2. The new vector, which forms part of the wider disclosure of the specification, was used in the cloning of the human CD2 antigen gene.

XX Sequence 2932 BP; 775 A; 698 C; 582 G; 777 T; 0 other;

Query Match		41.8%;	Score 755;	DB 13;	Length 2932;
Best Local Similarity		97.6%;	Pred. No. 3.3e-210;		
Matches 788;		Conservative 0;	Mismatches 15;	Indels 4;	Gaps 2;
QY	469	ATCGAATTCCTTCGGACTTTTGAAGTGTGCTGGGGAAGGATTCGAACCTTCGAA	528		
Db	804	AGCAGATTCCTTCGGACTTTTGAAGTGTGCTGGGGAAGGATTCGAACCTTCGAA	745		
QY	529	GTCGATGACGGCAGATTAGAGTCTCCTCTTGGCCGCTCGGAAACCCACACGGGT	588		
Db	744	GTCGATGACGGCAGATTAGAGTCTCCTCTTGGCCGCTCGGAAACCCACACGGGT	685		
QY	589	AATGCTTTTACTGGCTGCTCCCTTATCGGGAGCGGGCGCATCATCAATACGCG	648		
Db	584	AATGCTTTTACTGGCTGCTCCCTTATCGGGAGCGGGCGCATCATCAATACGCG	625		
QY	649	CCGCTGTAAGTGTAGCTTGTGAGAAGATTCCTCGAGCCCGCGCTTCTGGCGTTT	708		
Db	624	CCGCTGTAAGTGTAGCTTGTGAGAAGATTCCTCGAGCCCGCGCTTCTGGCGTTT	568		
QY	709	TCCATAGGCTCCGCCCTCGAGAGCATCACAAAATCGACCTCAAGTCAGAGTGGC	768		
Db	567	TCCATAGGCTCCGCCCTCGAGAGCATCACAAAATCGACCTCAAGTCAGAGTGGC	508		
QY	769	GAACCCGACAGGACTATAAGATACAGGCGTTTCCCTCGGAAGCTCCCTCGTGGCT	828		
Db	507	GAACCCGACAGGACTATAAGATACAGGCGTTTCCCTCGGAAGCTCCCTCGTGGCT	448		
QY	829	CTCCTGTTCCGACCTCGCGCTTACCGGATACCTGTGCGCTTCTCCCTCGGGAAGCG	888		
Db	447	CTCCTGTTCCGACCTCGCGCTTACCGGATACCTGTGCGCTTCTCCCTCGGGAAGCG	388		
QY	889	TGCGGCTTCTCAATGCTCAGCTGTAGTATCTCAGTTCGCTGTAGTTCGCTTCCA	948		
Db	387	TGCGGCTTCTCAATGCTCAGCTGTAGTATCTCAGTTCGCTGTAGTTCGCTTCCA	328		
QY	949	AGCTGGCTGTGTGACGAAACCCCGCTTACCGCGCGCTTATCCGCTTAACCT	1008		
Db	327	AGCTGGCTGTGTGACGAAACCCCGCTTACCGCGCGCTTATCCGCTTAACCT	268		
QY	1009	ATCGCTTTGAGTCCACCGGTAAGACAGCACTTATCGCCACTGCGCAGCACTGGTA	1068		
Db	267	ATCGCTTTGAGTCCACCGGTAAGACAGCACTTATCGCCACTGCGCAGCACTGGTA	208		
QY	1069	ACAGGATTACGAGGAGGTATGTAGGCGGTGCTACAGAGTCTTGAAGTGTGGCCTA	1128		
Db	207	ACAGGATTACGAGGAGGTATGTAGGCGGTGCTACAGAGTCTTGAAGTGTGGCCTA	148		
QY	1129	ACTAGGCTACACTAGAGGACAGTATTTGGTATCTGCGCTCTGCTGAAGCCAGTTACT	1188		
Db	147	ACTAGGCTACACTAGAGGACAGTATTTGGTATCTGCGCTCTGCTGAAGCCAGTT-CCT	89		
QY	1189	TCGAAAAAAGAGTTGGTAGCTCTTGATCCGGGAAACAAACCCCGCTGTAGCGTGGTT	1248		
Db	88	TCGAAAAAAGAGTTGGTAGCTCTTGATCCGGGAAACAAACCCCGCTGTAGCGTGGTT	29		

QY	1249	TTTTTGTTCGACGACGATTAACGC	1275	
Db	28	TTTTTGTTCGACGACGATTAACGC	2	
RESULT	5			
AAAT38557/c				
ID	AAAT38557	standard; DNA; 2932 BP.		
XX	AAAT38557;			
XX	21-JAN-1997	(first entry)		
XX	Expression vector	piH3.		
XX	CD19; cell surface antigen; T lymphocyte; immunoselection cloning;			
KW	vector; piH3; cyclic.			
XX	Synthetic.			
OS				
XX	Key	Location/Qualifiers		
FT	misc_feature	1..589		
FT		/*tag= a		
FT		/note= "bases derived from pMB1 origin (pBR322 ori)"		
FT	misc_feature	590..597		
FT		/*tag= b		
FT		/function= SacI linker		
FT	misc_feature	598..799		
FT		/*tag= c		
FT		/function= synthetic tyrosine suppressor tRNA gene		
FT	LTR	800..947		
FT		/*tag= d		
FT		/function= ASV LTR fragment (PvuII to MluI)		
FT	enhancer	948..1500		
FT		/*tag= e		
FT		/function= human cytomegalovirus AD169 enhancer		
FT	TATA_signal	1501..1650		
FT		/*tag= f		
FT		/function= HIV TATA and tat-responsive elements		
FT	misc_feature	1651..1761		
FT		/*tag= g		
FT		/function= pILNAN polylinker (HindIII to Xba)		
FT	misc_signal	1717..2569		
FT		/*tag= h		
FT		/function= pSV splice and poly-Addition signals		
FT	rep_origin	2570..2917		
FT		/*tag= i		
FT		/function= SV40 origin of replication (PvuII to HindIII)		
FT	misc_feature	2918..2922		
FT		/*tag= j		
FT		/function= piVX, remnant of RI site from polylinker		
XX	EP739980-A2.			
XX	30-OCT-1996.			
XX	23-FEB-1989;	89EP-0103127.		
XX	25-FEB-1988;	88US-0160416.		
XX	(MASS-) MASSACHUSETTS GEN HOSPITAL.			
XX	Allen J, Aruffo A, Camerini D, Lauffer L, Oquendo C;			
PI	Seed B, Simmons D, Stamenkovic I, Stengelin S;			
XX	WPI; 1996-478744/48.			
DR	cDNA encoding CD19 cell surface antigen - useful for recombinant			
XX	antigen prodn. for diagnostic and therapeutic purposes			
PT	Disclosure; Fig 1; 23pp; English.			
XX				

XX	Expression vector pIH3 (AAT38557), deposited as ATCC 67634, can be used in a novel immunoselection cloning method for cell surface antigen genes, e.g. human CD19 (see also AAT38556). This involves introducing cDNA fragments into the vector, transfecting mammalian cells in culture, maintaining the cells under conditions that allow antigen expression, exposing the cells to an antibody directed against the antigen, and recovering positive cells by contact with an immobilised antibody directed against the first antibody.									
XX	Sequence 2932 BP; 773 A; 700 C; 682 G; 777 T; 0 other;									
XX	Query Match 41.8%; Score 755; DB 17; Length 2932;									
XX	Best Local Similarity 97.6%; Pred. No. 3.3e-210;									
XX	Matches 788; Conservative 0; Mismatches 15; Indels 4; Gaps 2;									
QY	469	ATCGAATTTCTTCGGACTTTTGAAGTAGTGTGGGGGAAGGATTCGAACCTTCGAA	528							
Db	804	AGCAGATTCTTCGGACTTTTGAAGTAGTGTGGGGGAAGGATTCGAACCTTCGAA	745							
QY	529	GTGCGATGACGCGAGATTAGAGTCTCCCTTTGGCCGCTCGGGAACCCACCGGTT	598							
Db	744	GTGCGATGACGCGAGATTAGAGTCTCCCTTTGGCCGCTCGGGAACCCACCGGTT	685							
QY	589	AATGCTTTTACTGGCTGCTCCCTTATCGGGAAGCGGGCGCATCATCAATGACGCG	648							
Db	684	AATGCTTTTACTGGCTGCTCCCTTATCGGGAAGCGGGCGCATCATCAATGACGCG	625							
QY	649	CGGCTGTAAGTGTACGTTGAGAAGAAATCTCTGAGCCCGCGGTTGCTGGCGTTT	708							
Db	624	CGGCTGTAAGTGTACGTTGAGAAGAAATCTCTGAGCCCGCGGTTGCTGGCGTTT	568							
QY	709	TCATAGGCTCCGCCCTGAGGAGCATCAAAAATCGACGCTCAAGTCAGAGTGC	768							
Db	567	TCATAGGCTCCGCCCTGAGGAGCATCAAAAATCGACGCTCAAGTCAGAGTGC	508							
QY	769	GAACCCGACGAGACTATAAGATACAGCGGTTTCCCGCTGGAAGTCCCTCGTGGCT	828							
Db	507	GAACCCGACGAGACTATAAGATACAGCGGTTTCCCGCTGGAAGTCCCTCGTGGCT	448							
QY	829	CTCCTGTTCCGACCCGTCGCGTTACCGGATACGTCGCGCTTTCTCCCTTCGGGAAGCG	888							
Db	447	CTCCTGTTCCGACCCGTCGCGTTACCGGATACGTCGCGCTTTCTCCCTTCGGGAAGCG	388							
QY	889	TGCGGCTTTCTCAATGCTCACGCTGAGTATCTACGTTGCGGTGAGTTCGCTCCA	948							
Db	387	TGCGGCTTTCTCAATGCTCACGCTGAGTATCTACGTTGCGGTGAGTTCGCTCCA	328							
QY	949	AGCTGGGCTGTGTGCAACACCCCGTTTCAGCCCGACCGCTGCGCTTATCGGGTAACT	1008							
Db	327	AGCTGGGCTGTGTGCAACACCCCGTTTCAGCCCGACCGCTGCGCTTATCGGGTAACT	268							
QY	1009	ATCGCTTGAGTCCACACCGGTAAGACACACTTATCGCACTGGCAGCAGCACTGGTA	1068							
Db	267	ATCGCTTGAGTCCACACCGGTAAGACACACTTATCGCACTGGCAGCAGCACTGGTA	208							
QY	1069	ACAGGATTAGCAGCGAGGTATGTAGGCGGTCTACAGATTCTTGAAGTGTGGCTTA	1128							
Db	207	ACAGGATTAGCAGCGAGGTATGTAGGCGGTCTACAGATTCTTGAAGTGTGGCTTA	148							
QY	1129	ACTACGGCTACACTAGAAGACAGATTTGGTATCTGCGCTGCTGTAAGCCAGTTACCT	1188							
Db	147	ACTACGGCTACACTAGAAGACAGATTTGGTATCTGCGCTGCTGTAAGCCAGTT-CCT	89							
QY	1189	TCGGAATAAGAGTTGTAGTCTTTGATCCGGCAACAAACACCGCTGTAGCGTGGTT	1248							
Db	88	TCGGAATAAGAGTTGTAGTCTTTGATCCGGCAACAAACACCGCTGTAGCGTGGTT	29							
QY	1249	TTTTTCTTTGCAAGCAGCAGATTACGC	1275							
Db	28	TTTTTCTTTGCAAGCAGCAGATTACGC	2							

RESULT 6		
AAT14702/c		
ID	AAT14702 standard; DNA; 2932 BP.	
XX		
AC	AAT14702;	
XX		
DT	30-OCT-1996 (first entry)	
XX		
DE	Vector plasmid pIH3.	
XX		
KW	Cell surface antigen; cloning; immunoselection; immunotherapy;	
KW	therapy; diagnosis; vector; pIH3; CD2; T-cell receptor; COS;	
KW	T-lymphocyte; ss.	
XX		
OS	Synthetic.	
XX		
FH	Key	Location/Qualifiers
FT	misc_feature	1..589
FT	/*tag= a	
FT	/note= "bases 1-589 are derived from the pMB1 origin (pBR322 Ori)"	
FT		
FT	misc_feature	590..597
FT	/*tag= b	
FT	/note= "bases 590-597 are derived from the SacII linker"	
FT		
FT	misc_feature	598..799
FT	/*tag= c	
FT	/note= "bases 598-799 are derived from the synthetic tyrosine suppressor tRNA gene (supF gene)"	
FT		
FT	LTR	800..947
FT	/*tag= d	
FT	/note= "bases 800-947 are derived from a remnant of the ASV LTR fragment (PvuII-MluI)"	
FT		
FT	enhancer	948..1500
FT	/*tag= e	
FT	/note= "bases 948-1500 are derived from the hCMV AD169 enhancer"	
FT		
FT	misc_signal	1501..1650
FT	/*tag= f	
FT	/note= "bases 1501-1650 are derived from HIV TATA and tat-responsive elements"	
FT		
FT	misc_feature	1651..1716
FT	/*tag= g	
FT	/note= "bases 1651-1716 are derived from the pLXNAN polylinker (HindIII-Xba)"	
FT		
FT	misc_signal	1717..2569
FT	/*tag= h	
FT	/note= "bases 1717-2569 are derived from pSV to splice and poly-Addition signals"	
FT		
FT	misc_feature	2570..2917
FT	/*tag= i	
FT	/note= "bases 2570-2917 are derived from the SV40 origin of replication (PvuII-HindIII)"	
FT		
FT	misc_feature	2918..2922
FT	/*tag= j	
FT	/note= "bases 2918-2922 are derived from pIVX, remnant of R1 site from polylinker"	
FT		
XX	US5506126-A.	
XX		
PD	09-APR-1996.	
XX		
PF	25-FEB-1988;	88US-0160416.
XX		
PR	01-DEC-1992;	92US-0983647.
PR	25-FEB-1988;	88US-0160416.
PR	13-JUL-1989;	89US-0379076.
PR	18-JUL-1990;	90US-0553759.
PR	18-OCT-1993;	93US-0139273.
XX		
PA	(GEHO) GEN HOSPITAL CORP.	

XX	Aruffo A, Seed B;
PI	XX
XX	WPI; 1996-200279/20.
DR	XX
XX	Cloning of cDNA encoding cell surface antigen - useful for isolation
PT	of diagnostic and therapeutic proteins
PT	XX
XX	Example 1; Fig 1A-B; 79pp; English.
PS	XX
XX	Vector pIH3 (AA114702) is a COS cell expression vector constructed
CC	from piSV by inserting a synthetic transcription unit between the
CC	suppressor tRNA gene and the SV40 origin. This novel vector allows
CC	the generation of very large mammalian expression libraries and
CC	yields large amounts of protein in mammalian host cells, resulting
CC	in efficient selection. Efficiency of library construction is
CC	comparable to that achieved with phage lambda, but the clones
CC	generated are easier to manipulate. A novel immunoselection
CC	cloning method was used to clone genes (see also AA114703-04 and
CC	AA114706-26) for cell surface antigens of human lymphocytes.
XX	XX
SQ	Sequence 2932 BP; 774 A; 698 C; 683 G; 777 T; 0 other;
	Query Match 41.8%; Score 755; DB 17; Length 2932;
	Best Local Similarity 97.6%; Pred. No. 3.3e-210;
	Matches 788; Conservative 0; Mismatches 15; Indels 4; Gaps
QY	469 ATCGAATTCCTTCGGACHTTTCAAGTGATGCTGTGGGGGAAGATTTCGAACCTTCGAA 528
DB	804 AGCAGATTCTTTCGGACTTTTGAAGAAGTAGTGCTGTGGGGGAAGATTTCGAACCTTCGAA 745
QY	529 GTCCGATGACGGCAGATTTAGAGTCTGCTCCCTTTGGCCGCTCGGGAACCCACCACGGGT 588
DB	744 GTCCGATGACGGCAGATTTAGAGTCTGCTCCCTTTGGCCGCTCGGGAACCCACCACGGGT 685
QY	589 AATGCTTTTACTGGCTGCTCCCTTTATCGGAAGCGGGCGCATCATATCAAATGACGCG 648
DB	684 AATGCTTTTACTGGGCTGCTCCCTTTATCGGAAGCGGGCGCATCATATCAAATGACGCG 625
QY	649 CGCGCTAAAGTGTTACGTTTGAGAAGAATTCCTGCAGCCGCCGCTGTGCTGGCGTTTT 708
DB	624 CGCGCTAAAGTGTTACGTTTGAGAAGA ---CCGCGGTAATTCGGCGTTGTGGCGTTTT 568
QY	709 TCATAGGCTCGCCCCCTGACGAGCATCACAAAATCGAGCTCAAGTCAGAGGTGGC 768
DB	567 TCCATAGGCTCGCCCCCTGACGAGCATCACAAAATCGAGCTCAAGTCAGAGGTGGC 508
QY	769 GAACCCGACAGGACTATAAAGATACAGCGGTTTTCCCCTCGGAAGCTCCCTGTGCGCT 828
DB	507 GAACCCGACAGGACTATAAAGATACAGCGGTTTTCCCCTCGGAAGCTCCCTGTGCGCT 448
QY	829 CTCCTGTTCGACCCCTGCCGCTTACCGGATACCTGTCCGCTTTCTCCCTTCGGGAAGCG 888
DB	447 CTCCTGTTCGACCCCTGCCGCTTACCGGATACCTGTCCGCTTTCTCCCTTCGGGAAGCG 388
QY	889 TGGCGCTTTCTCAATGCTCACGCTGTAGGTATCTCAGTTTCGGTGTAGGTGCTTCGCTCCA 948
DB	387 TGGCGCTTTCTCATAGCTCACGCTGTAGGTATCTCAGTTTCGGTGTAGGTGCTTCGCTCCA 328
QY	949 AGCTGGGCTGTGTGCAGAACCCCGCTTCAGCCCGACACGCTCGCGCTTATCCGGTAACT 1008
DB	327 AGCTGGGCTGTGTGCAGAACCCCGCTTCAGCCCGACACGCTCGCGCTTATCCGGTAACT 268
QY	1009 ATCGCTTTGAGTCCAACCCGGTRAGACACGACTTATCGCCACTTGGCAGCAGCCACTGGTA 1068
DB	267 ATCGCTTTGAGTCCAACCCGGTRAGACACGACTTATCGCCACTTGGCAGCAGCCACTGGTA 208
QY	1069 ACAGGATTAGCAGAGCGAGGTATGTAGCGGTGCTACAGAGTTCTTGAATGTGTGGCCTA 1128
DB	207 ACAGGATTAGCAGAGCGAGGTATGTAGCGGTGCTACAGAGTTCTTGAATGTGTGGCCTA 148
QY	1129 ACTACGGCTACACTAGAAGGACATATTGGTATCTCGGCTCTCTCTGAAGCCAGTTACCT 1188

PR 13-JUL-1989; 89US-0379076.
 PR 13-JUL-1990; 90US-053759.
 PR 21-MAY-1997; 97US-0861205.
 XX (GEO) GEN HOSPITAL CORP.
 XX Aruffo A, Seed B;
 XX WPI; 1998-609251/51.
 XX New cloning vector and polylinker - based on existing sequences for
 PT efficient cloning and expression of mammalian cDNA(s), especially
 PT human lymphocyte antigenic sequences
 XX Example 1; Fig 1A-B; 75pp; English.
 XX This is the nucleotide sequence of expression vector piH3, a COS
 CC cell expression vector constructed from pISV by inserting a
 CC synthetic transcription unit between the suppressor tRNA gene and
 CC the SV40 origin. The vector can be used in a novel method for
 CC cloning cDNAs from mammalian expression libraries. The method is
 CC based on transient expression of an antigen in eukaryotic cells and
 CC immunoselection of cells expressing the antigen by adhesion to
 CC an antibody-coated substrate. The method is useful for the
 CC isolation and molecular cloning of any protein which can be
 CC expressed and transported to the cell surface membrane of a
 CC eukaryotic cell. It has been used to clone genes encoding cell
 CC surface antigens associated with mammalian T lymphocyte antigens
 CC such as CD1-53, LFA-3, ICAM-1, FCRI, TLISA and Leu8 (see AAV6342-63
 CC and AAV80440-55). The invention provides high efficiency expression
 CC vectors, such as piH3, which allow the generation of very large
 CC mammalian expression libraries, and yield large amounts of protein
 CC in mammalian host cells, resulting in efficient selection. The
 CC purified genes and proteins are useful for immunodiagnostic and
 CC immunotherapeutic applications, including the diagnosis and
 CC treatment of immune-mediated infections, diseases, and disorders of
 CC animals, including humans.
 XX Sequence 2932 BP; 774 A; 699 C; 682 G; 777 T; 0 other;

Query Match 41.8%; Score 755; DB 19; Length 2932;
 Best Local Similarity 97.6%; Pred. No. 3.3e-210;
 Matches 788; Conservative 0; Mismatches 15; Indels 4; Gaps 2;

QY 469 ATCGAATTTCTTCGGACTTTTGAAGTATGCTGGTGGGGAGGAGTTCGAACCTTCGAA 528
 DB 804 AGCAGATTTCTTCGGACTTTTGAAGTATGCTGGTGGGGAGGAGTTCGAACCTTCGAA 745
 QY 529 GTCGATGACGGCAGATTTAGAGTCTGCTCCTTTGGCCGCTCGGGAACCCACACGGGT 588
 DB 744 GTCGATGACGGCAGATTTAGAGTCTGCTCCTTTGGCCGCTCGGGAACCCACACGGGT 685
 QY 589 AATGCTTTTACTGGCTGCTCCCTTATCGGGAAGCGGGCGCATCATATCAAAATGACGG 648
 DB 684 AATGCTTTTACTGGCTGCTCCCTTATCGGGAAGCGGGCGCATCATATCAAAATGACGG 625
 QY 649 CCCTGTAAGTCTTACGTTGAGAAGAATTCCTGAGCCCGCGCTGCTGGCGTTT 708
 DB 624 CCCTGTAAGTCTTACGTTGAGAAGA--CCGCGGTAATTCGGCGTTGCTGGCGTTT 568
 QY 709 TCCATAGGCTCCGCCCTTCGAGCAGCATCACAAAATCGACGCTCAAGTCAGAGTGGC 768
 DB 567 TCCATAGGCTCCGCCCTTCGAGCAGCATCACAAAATCGACGCTCAAGTCAGAGTGGC 508
 QY 769 GAAACCCGACGAGCTATAAAGATACGAGCGTTTCCCTCGGAAGCTCCCTCGTGGCGT 828
 DB 507 GAAACCCGACGAGCTATAAAGATACGAGCGTTTCCCTCGGAAGCTCCCTCGTGGCGT 448
 QY 829 CTCCTGTTCCGACCCCTGCGGCTTACCGGATACCTGCTCGCGCTTTCTCCCTTCGGGAGCG 888
 DB 447 CTCCTGTTCCGACCCCTGCGGCTTACCGGATACCTGCTCGCGCTTTCTCCCTTCGGGAGCG 388

QY 889 TGGCGCTTTCTCAATGCTCACGCTGTAGGTATCTCAGTTTCGGTGTAGGTTCGCTCCA 948
 DB 387 TGGCGCTTTCTCATGCTCAGCGTGTAGGTATCTCAGTTTCGGTGTAGGTTCGCTCCA 328
 QY 949 AGCTGGGCTGTGTGACAGAACCCCGCTTCAGCCCGACCGCTCGGCTTATCCGGTAACT 1008
 DB 327 AGCTGGGCTGTGTGACAGAACCCCGCTTCAGCCCGACCGCTCGGCTTATCCGGTAACT 268
 QY 1009 ATGCTCTTGTAGTCCAAACCCGGTAAGACAGACTTATCGCCACTGGCAGCAGCCACTGGTA 1068
 DB 267 ATGCTCTTGTAGTCCAAACCCGGTAAGACAGACTTATCGCCACTGGCAGCAGCCACTGGTA 208
 QY 1069 ACAGGATTAGCAGAGGAGGTATGTAGGGGCTGTACAGAGTTCTTCAAGTGTGGCCTA 1128
 DB 207 ACAGGATTAGCAGAGGAGGTATGTAGGGGCTGTACAGAGTTCTTCAAGTGTGGCCTA 148
 QY 1129 ACTACGGCTACACTAGAAGACAGTATTTGGTATCTGCGGCTCTGTGAGCCAGTTACCT 1188
 DB 147 ACTACGGCTACACTAGAAGGACAGTATTTGGTATCTGCGGCTCTGTGAGCCAGTT-CCT 89
 QY 1189 TCGGAAAAAGAGTTGTAGTCTTGTATCGGCAACAAACCCGCTGTGTAGCGGTGGTT 1248
 DB 88 TCGGAAAAAGAGTTGTAGTCTTGTATCGGCAACAAACCCGCTGTGTAGCGGTGGTT 29
 QY 1249 TTTTGTGTTGCAAGCAGCAGATTACGC 1275
 DB 28 TTTTGTGTTGCAAGCAGCAGATTACGC 2
 RESULT 8
 AAV81199/C
 ID AAV81199 standard; DNA; 2932 BP.
 XX AAV81199;
 AC AAV81199;
 XX 10-MAY-1999 (first entry)
 DT Expression vector piH3.
 DE
 XX CD40; cell surface antigen; human; vector; plasmid piH3;
 KW cDNA library; COS; ss.
 XX Chimeric - Escherichia coli.
 OS Chimeric - Human cytomegalovirus.
 OS Chimeric - Human immunodeficiency virus.
 OS Chimeric - Rhesus macaque polyoma virus.
 XX Key Location/Qualifiers
 FT misc_feature 1..589
 FT /tag= a
 FT /note= "from pMB1 origin (pBR322 ori)"
 FT misc_feature 590..597
 FT /tag= b
 FT /note= "Sali linker"
 FT misc_feature 598..799
 FT /tag= c
 FT /note= "from synthetic tyrosine suppressor tRNA
 LTR 800..947
 FT /tag= d
 FT /note= "ASV LTR fragment"
 FT enhancer 948..1500
 FT /tag= e
 FT /note= "human cytomegalovirus ADL69 enhancer"
 FT TATA_signal 1501..1650
 FT /tag= f
 FT /note= "HIV TATA and tat-responsive elements"
 FT misc_feature 1651..1716
 FT /tag= g
 FT /note= "pLXNAN polylinker"
 FT misc_signal 1717..2569
 FT /tag= h
 FT /note= "SV40 small t antigen splice and early

Fri Jan 18 08:27:46 2002

FT misc_feature 2570..2917 region polyA signals from pSV2"
FT /*tag= i
FT /note= "SV40 origin of replication"
FT 2918..2922
FT /*tag= j
FT /note= "p1vX, remnant of RI site from polylinker"

XX US5849898-A.

PN 15-DEC-1998.

XX 07-JUN-1995; 95US-0485447.

XX 01-DEC-1992; 92US-0983647.

PR 25-FEB-1988; 88US-0160416.

PR 13-JUL-1989; 89US-0379076.

PR 23-MAR-1990; 90US-0498809.

PR 13-JUL-1990; 90US-0553759.

PR 07-JUN-1995; 95US-0485447.

XX (GEHO) GEN HOSPITAL CORP.

XX Allen J, Amlot M, Aruffo A, Camerini D, Lauffer L;

PI Oquendo C, Seed B, Simmons D, Stamenkovic I, Stengelin S;

XX WPI; 1999-069813/06.

XX cDNA encoding human CD40 antigen - useful for cloning cDNA encoding
PT cell surface antigens, constructing cDNA libraries, expression
PT vectors for expression in eukaryotic cells or their fragments

XX Example 1; Fig 1A-B; 79pp; English.

XX This is the nucleotide sequence of expression vector piH3, a COS
CC cell expression vector constructed from p1SV by inserting a
CC synthetic transcription unit between the suppressor tRNA gene and
CC the SV40 origin. The vector can be used in a novel method for
CC cloning cDNAs from mammalian expression libraries. The method is
CC based on transient expression of an antigen in eukaryotic cells and
CC physical selection of cells expressing the antigen by adhesion to
CC an antibody-coated substrate, such as a culture dish. The method
CC is useful for the isolation and molecular cloning of any protein
CC which can be expressed and transported to the cell surface membrane
CC of a eukaryotic cell. The invention specifically provides the CD40
CC cDNA sequence (see AAV81198). It also provides high efficiency
CC expression vectors, such as piH3, which allow the generation of
CC very large mammalian expression libraries, and yield large amounts
CC of protein in mammalian host cells, resulting in efficient selection.

XX SQ Sequence 2932 BP; 774 A; 699 C; 682 G; 777 T; 0 other;

Query Match 41.8%; Score 755; DB 20; Length 2932;
Best Local Similarity 97.6%; Pred. No. 3.3e-210;
Matches 788; Conservative 0; Mismatches 15; Indels 4; Gaps 2;

QY 469 ATCGAATCTTTTCGGACTTTTGAAGTGTGTTGGGGGAGGATTCGAACCTTCGAA 528

DB 804 AGCAGATTCTTCGGACATTTTGAAGTGTGTTGGGGGAGGATTCGAACCTTCGAA 745

QY 529 GTCGATGACGGCAGATTAGATCTGCTCCCTTTGGCGCTCGGGAACCCACCACGGGT 588

DB 744 GTCGATGACGGCAGATTAGATCTGCTCCCTTTGGCGCTCGGGAACCCACCACGGGT 685

QY 589 AATGCTTTTACTGGCTGCTCCCTTTATCGGGAAGCGGCGCATCATCAATGACGCG 648

DB 684 AATGCTTTTACTGGCTGCTCCCTTTATCGGGAAGCGGCGCATCATCAATGACGCG 625

QY 649 CCGCTGTAAGTGTACGTGTAAGAAATTCCTGACGCCCGCGGTGCTGGCGTTT 708

DB 624 CCGCTGTAAGTGTACGTGTAAGAAATTCCTGACGCCCGCGGTGCTGGCGTTT 568

QY 709 TCCATAGGCTCGCCCTCGCTGACGAGCATCAAAAATCGACGCTCAAGTCAGAGTGGC 768
DB 567 TCCATAGGCTCGCCCTCGCTGACGAGCATCAAAAATCGACGCTCAAGTCAGAGTGGC 508
QY 769 GAAACCCGACAGGACTATAAAGATACCGAGGCTTTCCCCCTGGAGCTCCCTCGTGGCCT 828
DB 507 GAAACCCGACAGGACTATAAAGATACCGAGGCTTTCCCCCTGGAGCTCCCTCGTGGCCT 448
QY 829 CTCCTGTTCCGACCTCGCGCTTACCGGATACCTGTCGCGCTTTCTCCCTTGGGAAAGCG 888
DB 447 CTCCTGTTCCGACCTCGCGCTTACCGGATACCTGTCGCGCTTTCTCCCTTGGGAAAGCG 388
QY 889 TGGCGCTTTCTCAATGCTCACGCTGTAGTATCTCAGTTTCGCTGTAGTTCGCTCCCA 948
DB 387 TGGCGCTTTCTCATAGCTACGCTGTAGTATCTCAGTTTCGCTGTAGTTCGCTCCCA 328
QY 949 AGCTGGCTGTGTGCACGAAACCCCGCTTACGCCCGACCGCTTATCCGGTAACT 1008
DB 327 AGCTGGCTGTGTGCACGAAACCCCGCTTACGCCCGACCGCTTATCCGGTAACT 268
QY 1009 ATGCTCTTGAGTCCAAACCCGCTTAAGACACGACTTATCGCCACTGGCAGCAGCTGGA 1068
DB 267 ATGCTCTTGAGTCCAAACCCGCTTAAGACACGACTTATCGCCACTGGCAGCAGCTGGA 208
QY 1069 ACAGGATACGAGCAGGATGTATGAGCGGTGCTACAGAGTTCTTGAAGTGTGGCGCTA 1128
DB 207 ACAGGATACGAGCAGGATGTATGAGCGGTGCTACAGAGTTCTTGAAGTGTGGCGCTA 148
QY 1129 ACTAGGCTACACTAGAGGACAGTATTTGGTATCTGCGCTGTGCTGAGCCAGTTACCT 1188
DB 147 ACTAGGCTACACTAGAGGACAGTATTTGGTATCTGCGCTGTGCTGAGCCAGTT-CCT 89
QY 1189 TCGGAAAAAGAGTGGTAGCTCTTGTATCCGGCAACCAACACCGCTGTAGCGGTGGTT 1248
DB 88 TCGGAAAAAGAGTGGTAGCTCTTGTATCCGGCAACCAACACCGCTGTAGCGGTGGTT 29
QY 1249 TTTTGTGTTGCAAGCAGCAGATTACGC 1275
DB 28 TTTTGTGTTGCAAGCAGCAGATTACGC 2
RESULT 9
AAA50577/c
ID AAA50577 standard; DNA; 2932 BP.
XX AAA50577;
XX 19-DEC-2000 (first entry)
XX Eukaryotic cell expression vector plasmid piH3.
XX Plasmid piH3; vector; eukaryotic cell; COS; cell surface antigen;
KW immunodiagnosis; diagnosis; immunotherapy; gene therapy;
KW immune disorder; infection; asthma; immune-complex disease;
KW amyloidosis; multiple sclerosis; parasitic disease;
KW immunoselection; panning; ds.
XX Chimeric - Rhesus macaque polyoma virus.
OS Chimeric - Escherichia coli.
OS Chimeric - Human cytomegalovirus.
OS Chimeric - Human immunodeficiency virus.
OS Chimeric - Synthetic.
XX Key Location/Qualifiers
FT misc_feature 1..589
FT /*tag= a
FT /function= "replication origin"
FT /note= "pBR322 ori from pMB1"
FT misc_feature 590..597
FT /*tag= b
FT /note= "derived from SacI linker"
FT misc_feature 598..799
FT /*tag= c

XX 29-AUG-2001 (first entry)

XX Expression vector pIH3 DNA used to express human lymphocyte antigens.

DE Human; lymphocyte cell surface antigen; immune-mediated disease; asthma;

DE infection; immune deficiency disorder; hypersensitivity; inflammation;

KW systemic lupus erythematosus; platelet disorder; rheumatoid arthritis;

KW transplant rejection; cyclic; circular; mutant; pIH3; ds.

XX Chimeric - Avian sarcoma virus.

OS Chimeric - Human cytomegalovirus.

OS Chimeric - Human immunodeficiency virus.

OS Chimeric - Rhesus macaque polyoma virus.

OS Synthetic.

XX Key Location/Qualifiers

PH misc_feature 1..589

FT /*tag= a

FT /note= "Derived from the pMB1 origin (pBR322 ori)"

FT misc_feature 590..597

FT /*tag= b

FT /note= "Derived from the SacII linker (ACGCGCT)"

FT misc_feature 598..799

FT /*tag= c

FT /note= "Derived from the synthetic tRNA gene (supF)"

FT misc_feature 800..947

FT /*tag= d

FT /note= "Derived from a remnant of the ASV LTR fragment (PvuII to MluI)"

FT enhancer 948..1500

FT /*tag= e

FT /note= "Derived from the human cytomegalovirus AD169 enhancer"

FT promoter 1501..1650

FT /*tag= f

FT /note= "Derived from the HIV TATA and tat-responsive elements"

FT misc_feature 1651..1716

FT /*tag= g

FT /note= "Derived from the pILNAN polylinker (HindIII to Xba)"

FT misc_signal 1717..2569

FT /*tag= h

FT /note= "Derived from the pSV2 splice and polyadenylation signals"

FT misc_feature 2570..2917

FT /*tag= i

FT /note= "Derived from the SV40 origin of replication (PvuII to HindIII)"

FT misc_feature 2918..2922

FT /*tag= j

FT /note= "Derived from pIVX, remnant of R1 site from polylinker"

XX US6218525-B1.

XX 17-APR-2001.

XX 01-DEC-1992; 92US-0983647.

XX 25-FEB-1988; 88US-0160416.

PR 13-JUL-1989; 89US-0379076.

PR 13-JUL-1990; 90US-0553759.

XX (GEHO) GEN HOSPITAL CORP.

XX Seed B, Aruffo A, Simmons D;

XX WPI; 2001-289848/30.

XX New recombinant DNA encoding CD28 useful for diagnosing and treating

XX immune-mediated diseases, infections or disorders, e.g. systemic lupus

erythematosus, asthma, transplant rejection, rheumatoid arthritis

Example 1; Fig 1A-1B; 72pp; English.

The present sequence for expression vector pIH3 is used to provide high level expression of human lymphocyte cell surface antigens in eukaryotic host cells. Another expression vector pIH3M (AAS03174) is also described. Various human lymphocyte cell surface antigen cDNA sequences (AAS03172, AAS03173, AAS03175-AAS03195) are described relating to the invention of a novel method of cloning cDNA encoding cell surface antigens and efficient construction of cDNA libraries. A genetically engineered cDNA sequence encoding the CD28 amino acid extracellular domain sequence (amino acids 1-134 given in AAU02437) and/or comprising nucleotides 100-759, 154-555 or 154-759 of the CD28 cDNA sequence (AAS03175) is also new. The purified genes and proteins are useful for immunodiagnostic and immunotherapeutic applications, such as in the diagnosis and treatment of immune-mediated diseases, infections or disorders in animals and humans. Such diseases include immune deficiency diseases, diseases of immediate type of hypersensitivity, asthma, hypersensitivity pneumonitis, systemic lupus erythematosus, rheumatoid arthritis, acute and chronic inflammation, platelet disorders, plasma and other cell neoplasms, parasitic diseases, multiple sclerosis, Guillain-Barre syndrome and tissue and organ transplant rejection. The sequences can also be used to identify, isolate and purify other antibodies and antigens.

Sequence 2932 BP; 774 A; 699 C; 682 G; 777 T; 0 other;

Query Match 41.8%; Score 755; DB 22; Length 2932;
Best Local Similarity 97.6%; Pred. NO. 3.e-210;
Matches 788; Conservative 0; Mismatches 15; Indels 4; Gaps 2;

QY 469 ATCGAATCTTTTCGGACTTTTGAAGTGTGTTGGGGGAGGATTTCGAACCTTCGAA 528
DB |||||
QY 529 GTGATGACGGCAGATTAGAGTCTCTCCCTTGGCCGCTCGGGAACCCACACGGGT 588
DB |||||
QY 744 GTCGATGACGGCAGATTAGAGTCTCTCCCTTGGCCGCTCGGGAACCCACACGGGT 685
QY 589 AATGCTTTTACTGGCTGCTCCCTTATCGGGAAGGGGGCGCATCATATCAATCAGCG 648
DB |||||
QY 684 AATGCTTTTACTGGCTGCTCCCTTATCGGGAAGGGGGCGCATCATATCAATCAGCG 625
QY 649 CCGCTGTAAGTGTACGTTGAGAAAGATTCTCGCAGCCCGCGTTCGTCGGGTGTTT 708
DB |||||
QY 624 CCGCTGTAAGTGTACGTTGAGAAAGATTCTCGCAGCCCGCGTTCGTCGGGTGTTT 568
QY 709 TCCATAGCTCCGCCGCCCTCGAGGAGCATCACAATAATCGACGCTCAAGTCAGAGTGCG 768
DB |||||
QY 567 TCCATAGCTCCGCCGCCCTCGAGGAGCATCACAATAATCGACGCTCAAGTCAGAGTGCG 508
QY 769 GAAACCCGACAGGACTATAAGATACAGAGCGTTTCCCTCGAAGCTCCCTCGTCGCT 828
DB |||||
QY 507 GAAACCCGACAGGACTATAAGATACAGAGCGTTTCCCTCGAAGCTCCCTCGTCGCT 448
QY 829 CTCCTGTTCCGACCTCGCGTTTACCGGATACCTGTCCGCTTTCTCCCTTCGGGAAGCG 888
DB |||||
QY 447 CTCCTGTTCCGACCTCGCGTTTACCGGATACCTGTCCGCTTTCTCCCTTCGGGAAGCG 388
QY 889 TGGCGCTTCTCAATGCTCAGCTGTAGGTATCTCAGTTCGGTGTAGGTGTTTCGGTCCA 948
DB |||||
QY 387 TGGCGCTTCTCATAGTCTACGCTGTAGGTATCTCAGTTCGGTGTAGGTGTTTCGGTCCA 328
QY 949 AGCTGGGCTGTGTCAGCAACCCCGTTTCAGCCCGAGCGCTTCGCGCTTATCCGTAAC 1008
DB |||||
QY 327 AGCTGGGCTGTGTCAGCAACCCCGTTTCAGCCCGAGCGCTTCGCGCTTATCCGTAAC 268
QY 1009 ATCGTCTTTCAGTCCAAACCCGGTAAGACAGCACTTATCCGCTGGCAGCCACTGGTA 1068
DB |||||
QY 267 ATCGTCTTTCAGTCCAAACCCGGTAAGACAGCACTTATCCGCTGGCAGCCACTGGTA 208

XX PR 23-AUG-1995; 95GB-0017263.
XX (CANC-) CANCER RES CAMPAIGN TECHNOLOGY.
XX PI Collins MKL, Cosset F, Takeuchi Y, Weiss RA;
XX WPI; 1997-179287/16.
XX Selectable retroviral packaging cell lines and expression constructs
PT - comprise selectable gene downstream of gene of interest, are
PT selectable due to the in-efficiency associated with translation
PT re-initiation
XX Claim 13; Fig 12; 79pp; English.
XX This sequence represents the recombinant expression plasmid FBdelPASAF.
CC This sequence is a packaging-deficient construct having a viral env gene
CC (in this case from pD114) and a selectable marker (SM). It is an example
CC of a recombinant expression vector (REV) of the invention, used to create
CC a packaging cell line. The REV's of the invention comprise a gene of
CC interest (GOI) and a SM gene. The SM gene is arranged downstream of the
CC GOI and a GOI associated stop codon is spaced from a start codon of the
CC SM gene to ensure that the SM protein is expressed as a result of
CC translation reinitiation. The cell lines are transformed with two REV's,
CC both are replication deficient, one contains the viral gag-pol gene, the
CC which are directly selectable and which provide for high expression of
CC the viral gene, high titre retroviral vectors may be obtained. The
CC packaging cell lines are useful for gene therapy. Prior packaging cell
CC lines using full length retroviral genomes as helper genomes were
CC isolated by cotransfecting them with plasmids encoding selectable
CC markers. However, the helper functions can be lost during the passages of
CC the cells in culture and the current packaging systems provide limited
CC titres of infectious retroviral vectors. Co-transfection with a plasmid
CC encoding a SM does not directly select the best gag-pol-env-expressing
CC cells. The new retroviral packaging cell lines overcome these problems.
XX Sequence 5865 BP; 1492 A; 1602 C; 1390 G; 1373 T; 8 other;
Query Match 33.4%; Score 604.4; DB 18; Length 5865;
Best Local Similarity 91.4%; Pred. No. 4.6e-166;
Matches 638; Conservative 0; Mismatches 60; Indels 0; Gaps 0;
QY 595 TTACTGCGCTGCTCCCTTATFCGGGAAGCGGGCGCATATATCAATGACGCGCGCTG 654
DB 3752 ttatcatgtctggatccagatctggcccatgcgccgcgagatcgtatnnnnacatgtgag 3811
QY 655 TAAAGTGTACGTTGAGAAAGAAATTCCTGTCAGCGCGCGCTGCTGCGCTTTTCCATA 714
DB 3812 caaaaggccagcaaaaggccaggaacccgtaaaaggccgctgtgctgttttccata 3871
QY 715 GGCTCGCGCCCGCTTACCGGATACCTGTGTCGCGCTTCTGCTGCGCTGCTGCTGCTGCTG 774
DB 3872 ggtcgcgccccctgaagcatcaaaaatcgacgctcaagtcadaggtg9cgaaacc 3931
QY 775 CGACGAGCTATAAGATACACGAGGCTTTCCTCGGAGCTCCCTCGTGCCTCTCTCTG 834
DB 3932 cgacagagactataagataccaggcgttttccctctggaagctcctcgtgctctcgt 3991
QY 835 TTCCGACCTGCGCTTACCGGATACCTGTGTCGCGCTTCTGCTGCGGAGCGTGGCGG 894
DB 3992 ttccgacctgcgcttacggatacctgtccgcttttctcctcctcgggagcg9g9cg 4051
QY 895 TTTCTCAATGCTACCGCTAGGTATCTCAGTTCGCTGTAGTTCGCTGCTCCAAAGCTGG 954
DB 4052 ttctcaatgctacagctgtagtatactcagttcgtgtagtgctgctccaagctgg 4111
QY 955 GCTGTGTGACGAAACCCCGCTTACGCCGAGCGCTGCGCTTATCCGGTACTATCTGTC 1014
DB 4112 gctgtgtgacgaacccccctttcagcccgaccgctgacctatcccgtaactatcgtc 4171

QY 1015 TTAGTCCAAACCCCGTAAAGCACGACTTATCGCCACTGGCAGCAGCCACTGGTAACAGGA 1074
DB 4172 ttgagtccaacccggttaagacacgacttatcgccactggcagcagccactggtaacagga 4231
QY 1075 TTAGCAGAGCGAGGTATGAGCGGTGCTACAGAGTTCCTTGAAGTGGTGGCTTAACCTACG 1134
DB 4232 ttgacagagcaggtatgtagcggtgctacagagttcttgaagtgggtgcttaactcacg 4291
QY 1135 GGTACACTAGAGGACAGTATTTGCTGCTGCTGCTGAAGCCAGTTCCTTACCTTCGGAA 1194
DB 4292 gctacactagaaggacagattttggatctctgctctctgctgtaagccaggttaccttcgaa 4351
QY 1195 AAGAGTGTGGTGTAGCTTGTGATCCGCGCAACAAACACCGCTGCTGAGCGTGGTGGTGGTGGTGG 1254
DB 4352 aaagagttggtgagcttctgctgctgcaaaacacacacgctggtgagcggtggtttttg 4411
QY 1255 TTGCAAGCAGCAGATTACCGCGCAGAAAAAAGAGATCT 1292
DB 4412 ttgcaagcagcagattacgcgcagaaaaaaggatct 4449
RESULT 13
AAT90689
ID AAT90689 standard; DNA; 6028 BP.
XX AAT90689;
XX AC
XX AC
XX 05-JAN-1998 (first entry)
XX Plasmid FBdelPASAF coding sequence.
XX Packaging-deficient construct; viral gag-pol gene; packaging cell line;
XX moloney murine leukaemia virus; MoMLV; viral env gene; helper construct;
XX gene therapy; ss.
XX Synthetic.
XX OS
XX WO9708330-A1.
XX PD
XX 06-MAR-1997.
XX 23-AUG-1996; 96WO-GB02061.
XX 23-AUG-1995; 95GB-0017263.
XX (CANC-) CANCER RES CAMPAIGN TECHNOLOGY.
XX Collins MKL, Cosset F, Takeuchi Y, Weiss RA;
XX WPI; 1997-179287/16.
XX Selectable retroviral packaging cell lines and expression constructs
PT - comprise selectable gene downstream of gene of interest, are
PT selectable due to the in-efficiency associated with translation
PT re-initiation
XX Claim 13; Fig 9; 79pp; English.
XX This sequence represents the recombinant expression plasmid FBdelPASAF.
CC This sequence is a packaging-deficient construct having a viral env gene
CC (in this case from the moloney murine leukaemia virus) and a selectable
CC marker (SM). It is an example of a recombinant expression vector (REV) of
CC the invention, used to create a packaging cell line. The REV's of the
CC invention comprise a gene of interest (GOI) and a SM gene. The SM gene is
CC arranged downstream of the GOI and a GOI associated stop codon is spaced
CC from a start codon of the SM gene to ensure that the SM protein is
CC expressed as a result of translation reinitiation. The cell lines are
CC transformed with two REV's, both are replication deficient, one contains
CC the viral gag-pol gene, the other the viral env gene. By using helper
CC constructs, such as the REV's, which are directly selectable and which
CC provide for high expression of the viral gene, high titre retroviral
CC vectors may be obtained. The packaging cell lines are useful for gene
CC therapy. Prior packaging cell lines using full length retroviral genomes

CC as helper genomes were isolated by cotransfecting them with plasmids
CC encoding selectable markers. However, the helper functions can be lost
CC during the passages of the cells in culture and the current packaging
CC systems provide limited titres of infectious retroviral vectors.
CC Co-transfection with a plasmid encoding a SM does not directly select the
CC best gag-pol-env-expressing cells. The new retroviral packaging cell
CC lines overcome these problems.

XX SQ Sequence 6028 BP; 1515 A; 1637 C; 1499 G; 1369 T; 8 other;

Query Match 33.4%; Score 604.4; DB 18; Length 6028;
Best Local Similarity 91.4%; Pred. No. 4.7e-166;
Matches 638; Conservative 0; Mismatches 60; Indels 0; Gaps 0;

QY 595 TTTACTGGCTGCTCCCTTATCGGAAGCGGCGCATCATATCAATACGCGCCGCTG 654
Db 3915 ttatcatgtctggtccagatctg99cccatcg99ccgctcgatcnnnnacatgtgag 3974
QY 655 TAAAGTGTACGTTGAGAAAGAAATTCCTGCAGCCCGCGGCTTCTGGCGTTTTCCTATA 714
Db 3975 caaaagccagcaaaagccaggaacgtaaaagccgctgtgctgtttttccata 4034
QY 715 GGCTCCGCCCGCTGAGCAGCATCACAAAATCGACGCTCAAGTCAGAGTGGCGAAACC 774
Db 4035 ggtccgccccctgacgagcattcacaaaaatcgactcaagtcagaggtg99gaaacc 4094
QY 775 CGACGAGCTATAAAGATACAGCGGTTTCCCTCGTGAAGCTCCCTGCTGCTCTCTG 834
Db 4095 cgacagactataaagatacagcggttcccccctggaagctccctgtgctctctg 4154
QY 835 TTCGAGCCCTGCGCTTACCGGATACGTCGCGCTTCTCCCTTCGCGAAGCGTGGCGC 894
Db 4155 ttcgacctgcgcttacgatacctgtccgcttttcccttcggaagcggtg99c 4214
QY 895 TTTCTCAATGCTCAGCGTGTAGGTATCTAGTTCGGTGTAGGTGCTTCCAAAGCTGG 954
Db 4215 ttctcaatgctcaagctgtaggtatctcagttcgggtgtaggtcgttcgctccaagctg 4274
QY 955 GCTGTGTGACGAACCCCGCTTCAGCCCGACGCTGCGCTTATCCGTTAACTATCGTC 1014
Db 4275 gctgtgtgacgaacccccgcttcagcccgctgctgcttccggttaactatcgtc 4334
QY 1015 TTGAGTCCAAACCCGGTAAGACACACTTATCGCACCTGGCAGCAGCCACTGGTAACAGGA 1074
Db 4335 ttgagtcacacccggttaagacagacttatcgccactggcagcagcactggtaacagga 4394
QY 1075 TTAGCAGAGGAGTATGTAGCGCGTCTACAGAGTTCTTGAAGTGTGGCTTAACCTACG 1134
Db 4395 ttagcagagcaggtatgtaggcgtgctacagagttcttgaaagtgggtg99cctaactacg 4454
QY 1135 GCTACACTAGAAGACAGATTTTGGTATCTGCGCTGCTCAAGCCAGTTACCTTCGGAA 1194
Db 4455 gctacactagaagacagattttggtatcgcgtctgctgctgaagccaggttacctcggaa 4514
QY 1195 AAAGAGTTGGTGTAGCTTGTATCGCGCAAAACAAACACCGCTGTGAGCGGTGTTTTTGG 1254
Db 4515 aaagagttggtatgctctgtatccg99caaaacacccacctggtgtagcgtgtttttt 4574
QY 1255 TTTGCAAGCAGCAGATTTACGCGCAGAAAAAAGGATCT 1292
Db 4575 ttgcaagcagcagattacgcgcagaaaaaaggatct 4612

RESULT 14
AAT90690

ID AAT90690 standard; DNA: 6061 BP.

XX AC AAT90690;

XX DT 05-JAN-1998 (first entry)

XX XX Plasmid FBDelPMOSAF coding sequence.

XX Packaging-deficient construct; viral gag-pol gene; packaging cell line;
KW moloney murine leukaemia virus; MOMLV; viral env gene; helper construct;
KW gene therapy; ss.
XX Synthetic.

OS WO9708330-A1.

PN 06-MAR-1997.

XX 23-AUG-1996; 96WO-GB02061.

XX 23-AUG-1995; 95GB-0017263.

XX (CANC-) CANCER RES CAMPAIGN TECHNOLOGY.

XX Collins MKL, Cosset F, Takeuchi Y, Weiss RA;

PI WPI; 1997-179287/16.

XX Selectable retroviral packaging cell lines and expression constructs

PT - comprise selectable gene downstream of gene of interest, are

PT selectable due to the in-efficiency associated with translation

PT re-initiation

XX Claim 13; Fig 10; 79pp; English.

XX This sequence represents the recombinant expression plasmid FBDelPMOSAF.

CC This sequence is a packaging-deficient construct having a viral env gene

CC in this case from the moloney murine leukaemia virus) and a selectable

CC marker (SM). It is an example of a recombinant expression vector (REV) of

CC the invention, used to create a packaging cell line. The REV's of the

CC invention comprise a gene of interest (GOI) and a SM gene. The SM gene is

CC arranged downstream of the GOI and a GOI associated stop codon is spaced

CC from a start codon of the SM gene to ensure that the SM protein is

CC expressed as a result of translation reinitiation. The cell lines are

CC transformed with two REV's, both are replication deficient, one contains

CC the viral gag-pol gene, the other the viral env gene. By using helper

CC constructs, such as the REV's, which are directly selectable and which

CC provide for high expression of the viral gene, high titre retroviral

CC vectors may be obtained. The packaging cell lines are useful for gene

CC therapy. Prior packaging cell lines using full length retroviral genomes

CC as helper genomes were isolated by cotransfecting them with plasmids

CC encoding selectable markers. However, the helper functions can be lost

CC during the passages of the cells in culture and the current packaging

CC systems provide limited titres of infectious retroviral vectors.

CC Co-transfection with a plasmid encoding a SM does not directly select the

CC best gag-pol-env-expressing cells. The new retroviral packaging cell

CC lines overcome these problems.

XX SQ Sequence 6061 BP; 1498 A; 1665 C; 1485 G; 1405 T; 8 other;

Query Match 33.4%; Score 604.4; DB 18; Length 6061;

Best Local Similarity 91.4%; Pred. No. 4.7e-166;

Matches 638; Conservative 0; Mismatches 60; Indels 0; Gaps 0;

QY 595 TTTACTGGCTGCTCCCTTATCGGAAGCGGCGCATCATATCAATACGCGCCGCTG 654

Db 3948 ttatcatgtctggtccagatctg99cccatcg99ccgctcgatcnnnnacatgtgag 4007

QY 655 TAAAGTGTACGTTGAGAAAGAAATTCCTGCAGCCCGCGGCTTCTGGCGTTTTCCTATA 714

Db 4008 caaaagccagcaaaagccaggaacccgtaaaagccgctgtgctgtttttccata 4067

QY 715 GGCTCCGCCCGCTGAGCAGCATCACAAAATCGACGCTCAAGTCAGAGTGGCGAAACC 774

Db 4068 ggtccgccccctgacgagcattcacaaaaatcgactcaagtcagaggtg99gaaacc 4127

QY 775 CGACGAGCTATAAAGATACAGCGGTTTCCCTCGTGAAGCTCCCTGCTGCTCTCTG 834

Db 4128 cgacagactataaagatacagcggtttcccccctggaagctccctgctgctcctg 4187

(in this case from GALV) and a selectable marker (SM). It is an example of a recombinant expression vector (REV) of the invention, used to create a packaging cell line. The REV's of the invention comprise a gene of interest (GOI) and a SM gene. The SM gene is arranged downstream of the GOI and a GOI associated stop codon is spaced from a start codon of the SM gene to ensure that the SM protein is expressed as a result of translation reinitiation. The cell lines are transformed with two REV's, both are replication deficient, one contains the viral gag-pol gene, the other the viral env gene. By using helper constructs, such as the REV's, which are directly selectable and which provide for high expression of the viral gene, high titre retroviral vectors may be obtained. The packaging cell lines are useful for gene therapy. Prior packaging cell lines using full length retroviral genomes as helper genomes were isolated by cotransfecting them with plasmids encoding selectable markers. However, the helper functions can be lost during the passages of the cells in culture and the current packaging systems provide limited titres of infectious retroviral vectors. Co-transfection with a plasmid encoding a SM does not directly select the best gag-pol-env-expressing cells. The new retroviral packaging cell lines overcome these problems.

```
Query Match          33.4%; Score 604.4; DB 18; Length 6312;
Best Local Similarity 91.4%; Pred. No. 4.8e-166;
Matches 638: Conservative 0; Mismatches 60; Indels 0; Gaps 0;
```

[illegible]

QY	835	TTTCGACCCCTGCCGCTTACCGGATACCTGTCCGCCCTTCTCCCTTCGGGAAGCGTGGCGC	894
Db	4188	ttccgaccctgcgcttacccggataacctgcccgccttctcccttcggaaagctggcgc	4247
QY	895	TTTCTCAATGCTCAGCGTGTAGGTATCTCAGTTTCGGTGTAGTGTTCGCTCCAGCTGG	954
Db	4248	tttccaatgctcagctgtagttatctcaagttccggttagtctgcgtcccaagctgg	4307
QY	955	GCTGTGTGCAGAACCCCGCTTCAGCCCGACCGCTGCGCCTTATCCCGGTAACATATCGTC	1014
Db	4308	gctgtgtgcagaaaccccccgttcagcccagccgctgcgccttatccgtaactatgctc	4367
QY	1015	TTGAGTCCAAACCGGTGAAGACACGACTTATTCGCCACTTGGCAGCGCCACTGGTAACAGGA	1074
Db	4368	ttgagttcaaacccgtaagacaagacttatcgccactggcagcgcactggttaacagga	4427
QY	1075	TTAGCAGACCGAGGTATGTAGCCGGTGCTACAGAGTTCTTGAAGTGTGGCGCTAACTACG	1134
Db	4428	ttagcagagcggaggtatgtaggcgggtgtacagagttcttgaagtgtggcctaactacg	4487
QY	1135	GCTACACTAGAAGGACAGTATTTGTTATCTCGCTCTGCTGAAGCCAGTTACCTTCGGAA	1194
Db	4488	gctactactagaagacagattttggtatctgcctctgctgaagccaggttaacctcggaa	4547
QY	1195	AAGAGTTGGTAGCTCTTGATCCCGCAAAACACACCGCTGGTAGCGGTGGTTTTTTG	1254
Db	4548	aaagagttgtagctctgtatccgggggcaaacacacacgcgtggtagcgggtgttttttg	4607
QY	1255	TTTGCAGCAGCAGATATACGCGCAGAAAAAAGGATCT	1292
Db	4608	tttgcagcagcagatttacgcgagaaaaaaggatct	4645

RESULT 15

AAAT90691	AAAT90691 standard; DNA; 6312 BP.
XX	AC
XX	AAAT90691;
XX	05-JAN-1998 (first entry)
XX	Plasmid FdelPGASAF coding sequence.
XX	Packaging-deficient construct; viral gag-pol gene; packaging cell line;
KW	moloney murine leukaemia virus; MOMLV; viral env gene; helper construct;
KW	gene therapy; ss.
XX	Synthetic.
XX	WO9708330-A1.
XX	06-MAR-1997.
XX	23-AUG-1996; 96WO-GB02061.
XX	23-AUG-1995; 95GB-0017263.
XX	(CANC-) CANCER RES CAMPAIGN TECHNOLOGY.
XX	Collins MKL, Cosset F, Takeuchi Y, Weiss RA;
XX	WPI; 1997-179287/16.
XX	Selectable retroviral packaging cell lines and expression constructs
PT	- comprise selectable gene downstream of gene of interest; are
PT	selectable due to the in-efficiency associated with translation
PT	re-initiation
XX	Claim 13; Fig 11; 79pp; English.
XX	This sequence represents the recombinant expression plasmid FdelPASAF.
CC	This sequence is a packaging-deficient construct having a viral env gene

Db 4859 ttgcaagcagcagattacgcgcagaaaaaggatct 4896

Search completed: January 17, 2002, 12:05:08
Job time: 17895 sec

THIS PAGE BLANK (USPTO)

FEATURE:
NAME/KEY: misc_feature
LOCATION: (3801)
OTHER INFORMATION: n is any nucleotide
FEATURE:
NAME/KEY: misc_feature
LOCATION: (3802)
OTHER INFORMATION: n is any nucleotide
US-09-011-745-8

Query Match 33.4%; Score 604.4; DB 4; Length 5865;
Best Local Similarity 91.4%; Pred. No. 8.8e-175;
Matches 638; Conservative 0; Mismatches 60; Indels 0; Gaps 0;
QY 595 TTTACTGGCCCTGCTCCCTTATTCGGGAAGCGGGGCGCATCATCAAAATGACGCCCGCTG 654
Db 3752 ttatcatgtctggatccagatctggcccatcgccgcgcatgatnnnnacatgtgag 3811
QY 655 TAAAGTGTACGTTGAGAAAGAAATTCCTGAGCCCGGCTGCTGGCGTTTTCCTG 714
Db 3812 caaagccagcaaaagccaggaacccgtaaaagccgctgtgctggttttccata 3871
QY 715 GGCTCCGCCCGCCCTGAGGAGCATCACAAAATCGACGCTCAAGTCAGAGTGGCGAAAC 774
Db 3872 ggtccgccccctgacagcatcacaaaatcgacgctcaagtcagagtgcgaaacc 3931
QY 775 GCAGAGACTATAAGATACAGCGGTTTCCCGCTGGAAGCTCCCTGCTGCGTCTCCTG 834
Db 3932 cgacagactataaagataacagcggttcccccgtgaagctccctgctgctctctg 3991
QY 835 TTCGACCCCTGCGCTTACCGGATACCTGTCGCGCTTTCCTCCCTGCGGAAGCGTGGCG 894
Db 3992 ttcgaccccgccgcttacccgatactgctccgcttcccttcccttccggaagcgtggcg 4051
QY 895 TTTCTCAATGCTCACGCTGAGGTATCTCAGTTCGGGTAGTGTGCTGCTCCCAAGCTGG 954
Db 4052 ttctcaatgctcaagctgtagtatctcaagttcggtgtaggttcgctcccaagctgg 4111
QY 955 GCTGTGTGACGAAACCCCGCTTACGCGGACCGCTGCGCTTATCCGGTAACTATCGTC 1014
Db 4112 gctgtgtgcaagcaaccccgcttcagccgacccgctgcttcccttcccttccggaagcgtggcg 4171
QY 1015 TTGAGTCCAAACCGGTAAGACAGACTTATCGCCACTGCGCAGCAGCCACTGTAAACAG 1074
Db 4172 ttgagtcacacccggttaagacagacttatcgccactggcagcagcactggttaacagga 4231
QY 1075 TTAGCAGAGCGAGGTATGAGCGGTGCTACAGAGTTCCTGAAGTGGTGGCTTAACAG 1134
Db 4232 ttgacagagcaggtatgtagcggtgctacagagttcttgagtggtggtggttaactacg 4291
QY 1135 GCTACACTAAGAGGACAGTATTTGGTATCTGCGCTGCTGCTGAGCCAGTACCTTCGGAA 1194
Db 4292 gctacactagaagagcagatttggatctgctgctgctgctgagccaggttaacttcggaa 4351
QY 1195 AAAGAGTTGGTAGCTCTTATCCGGCAAAACCAACACCGCTGCTAGCGGTGTTTTTGG 1254
Db 4352 aaagattgtagctctgtagtcggtccgcaaacacacccgctggtgagcgtgttttttgg 4411
QY 1255 TTGCAAGCAGCAGATTACGCGCAGAAAGAGATCT 1292
Db 4412 ttgcaagcagcagattacgcgacagaaaaaaggatct 4449

RESULT 2

US-09-011-745-5
Sequence 5, Application US/09011745
Patent No. 6165715
GENERAL INFORMATION:
APPLICANT: Collins, Mary KL
APPLICANT: Weiss, Robin A
APPLICANT: Takeuchi, Yasuhiro
APPLICANT: Cosset, Francois-Loic

TITLE OF INVENTION: Expression systems
FILE REFERENCE: 09/011,745
CURRENT APPLICATION NUMBER: US/09/011,745
CURRENT FILING DATE: 1998-06-22
EARLIER APPLICATION NUMBER: PCT/GB96/02061
EARLIER FILING DATE: 1996-08-23
EARLIER APPLICATION NUMBER: GB9517263.1
EARLIER FILING DATE: 1995-08-23
NUMBER OF SEQ ID NOS: 29
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 5
LENGTH: 6028
TYPE: DNA
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: Description of Artificial Sequence: Portion of
OTHER INFORMATION: construct
FEATURE:
NAME/KEY: misc_feature
LOCATION: (3774)
OTHER INFORMATION: n is any nucleotide
FEATURE:
NAME/KEY: misc_feature
LOCATION: (3775)
OTHER INFORMATION: n is any nucleotide
FEATURE:
NAME/KEY: misc_feature
LOCATION: (3776)
OTHER INFORMATION: n is any nucleotide
FEATURE:
NAME/KEY: misc_feature
LOCATION: (3777)
OTHER INFORMATION: n is any nucleotide
FEATURE:
NAME/KEY: misc_feature
LOCATION: (3962)
OTHER INFORMATION: n is any nucleotide
FEATURE:
NAME/KEY: misc_feature
LOCATION: (3963)
OTHER INFORMATION: n is any nucleotide
FEATURE:
NAME/KEY: misc_feature
LOCATION: (3964)
OTHER INFORMATION: n is any nucleotide
FEATURE:
NAME/KEY: misc_feature
LOCATION: (3965)
OTHER INFORMATION: n is any nucleotide
US-09-011-745-5

Query Match 33.4%; Score 604.4; DB 4; Length 6028;
Best Local Similarity 91.4%; Pred. No. 8.9e-175;
Matches 638; Conservative 0; Mismatches 60; Indels 0; Gaps 0;
QY 595 TTTACTGGCCCTGCTCCCTTATTCGGGAAGCGGGGCGCATCATCAAAATGACGCCCGCTG 654
Db 3915 ttatcatgtctggatccagatctggcccatcgccgcgcatgatnnnnacatgtgag 3974
QY 655 TAAAGTGTACGTTGAGAAAGAAATTCCTGAGCCCGGCTGCTGCGGTTCCTCCATA 714
Db 3975 caaaagcagcaaaagccaggaacccgtaaaagccgctgtgctggttttccata 4034
QY 715 GGCTCCGCCCGCCCTGAGGAGCATCACAAAATCGACGCTCAAGTCAGAGTGGCGAAAC 774
Db 4035 ggtccgccccctgacagcatcacaaaatcgacgctcaagtcagagtggtgcaaac 4094
QY 775 GCAGAGACTATAAGATACAGCGGTTTCCCGCTGGAAGCTCCCTGCTGCTCTCCTG 834
Db 4095 cgacagactataaagataccagcggtttcccccgtgaagctccctgctgctctctg 4154
QY 835 TTCGACCCCTGCGCTTACCGGATACCTGTCCCGCTTTCCTTCGGAAGCGTGGCGC 894

```

|||||
Db 4155 ttccgaccctgcgcttacccgataccctgtccgcctttcccttccttcggaagcgtggcg 4214
Qy 895 TTTCTCAATGCTCAGCTGTAGGTATCTCAGTTTCGGTGTAGGTGTTCGCTCCAGCTGG 954
Db 4215 ttctcaatgctcacgctgtaggatctcagttcagttcgggtgaggtgttcgctccaagctgg 4274
Qy 955 GCTGTGTGCAGAACCCCGCTTCAGCCGACCGCTGCGCTTATCCGGTAACTATCGTC 1014
Db 4275 gctgtgtgcagcaaccctccttcagcccgaccctgcgcttaccggttaactatcgtc 4334
Qy 1015 TTGAGTCCAAACCCGGTAAGACACGACTTATGCCACTGGCAGCAGCCACTGGTAAACAGGA 1074
Db 4335 ttgagtcacaaccggtaagacacgacttatcgccactggcagcagcactggttaacagga 4394
Qy 1075 TTACGAGCGAGCTATCTAGCGGTGCTACAGAGTTCTTGAAGTGTGGCTTAACACTACG 1134
Db 4395 ttgacagcgaggtatgtaggcgtgctacagagttcttgaagtgtggcctaactacg 4454
Qy 1135 GCTACACTAGAAAGACAGTATTTGGTATCTGCGTCTGCTGAAGCCAGTTACCTTCGGAA 1194
Db 4455 gctacactagaagacagatttgttatctgtcgtctgtgaagccagttaccttcggaa 4514
Qy 1195 AAAGAGTGTGTAGTCTTGTATCCGGCAAAACAAACACCGCTGGTAGCGGTGGTTTTTGTG 1254
Db 4515 aaagagttggtagctcttgatcccggaacaaacacccgctggtagcggtggtttttttg 4574
Qy 1255 TTTCGAACGACAGATACGCGCGCAGAAAAAGGATCT 1292
Db 4575 ttgcaagcagcagattacgcgcgcaaaaaaaggatct 4612
```

```

RESULT 3
US-09-011-745-6
; Sequence 6, Application US/09011745
; Patent No. 6165715
; GENERAL INFORMATION:
; APPLICANT: Collins, Mary KL
; APPLICANT: Weiss, Robin A
; APPLICANT: Takeuchi, Yasuhiro
; APPLICANT: Cosset, Francois-Loic
; TITLE OF INVENTION: Expression systems
; FILE REFERENCE: 09/011,745
; CURRENT APPLICATION NUMBER: US/09/011,745
; EARLIER FILING DATE: 1998-06-22
; EARLIER FILING DATE: 1996-08-23
; EARLIER APPLICATION NUMBER: GB9517263.1
; EARLIER FILING DATE: 1995-08-23
; NUMBER OF SEQ ID NOS: 29
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 6
; LENGTH: 6061
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: Portion of
; FEATURE:
; OTHER INFORMATION: construct
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (3807)
; OTHER INFORMATION: n is any nucleotide
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (3808)
; OTHER INFORMATION: n is any nucleotide
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (3809)
; OTHER INFORMATION: n is any nucleotide
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (3810)
```

```

; OTHER INFORMATION: n is any nucleotide
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (3995)
; OTHER INFORMATION: n is any nucleotide
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (3996)
; OTHER INFORMATION: n is any nucleotide
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (3997)
; OTHER INFORMATION: n is any nucleotide
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (3998)
; OTHER INFORMATION: n is any nucleotide
; US-09-011-745-6

Query Match      33.4%: Score 604.4; DB 4; Length 6061;
Best Local Similarity 91.4%; Pred. No. 8.9e-175;
Matches 636; Conservative 0; Mismatches 60; Indels 0; Gaps 0;

Qy 595 TTTACTGCGCTGCTCCCTTATCGGGAGCGGGGCGCATCATATCAAAATGACGCGCGCTG 654
Db 3948 ttatcagtcgtggtacccagatctggcccatgcgcccggatcgatnnnnacatgtgag 4007
Qy 655 TAAAGTGTGTAGTGTGAGAAATTCCTGTGAGCGCGCGCTTGTGCGGTGTTCCTATA 714
Db 4008 caaagccagcaaaagccaggaaacgtaaaaagccgctgtgctgctgtttttccata 4067
Qy 715 GCTCCGCGCCCTCAGCAGCATCACAAAAATCGACGCTCAAGTCAGAGGTGGCGAAACC 774
Db 4068 ggcctccgccccctgacgagcatcaaaaaatcgacgtcaagtcaaggtggcgaaacc 4127
Qy 775 CGACAGGACTATAAAGATACACAGCGCTTTCGCCCTGGAAGCTCCCTCGTGGCGCTCTCCTG 834
Db 4128 cgacagactataaagataaccagcggtttcccccctggaagctccctcgtgcgtctcctg 4187
Qy 835 TTCGACCGCTGCGCTTACCGGATACCTGTCCGCTTTCCTCCCTTCGGGGAAGCGTGGCGC 894
Db 4188 ttccgacctgcgcttaccggatcacctgtccgccttctcccttcgggaagcgtggcg 4247
Qy 895 TTTCTCAATGCTCAGCTGTAGGTATCTCAGTTCGGTGTAGGTGCTGCTCCAGCTGG 954
Db 4248 ttctcaatgctcacgctgtgtaggtatctcagttcgtggtgctgctcccaagctgg 4307
Qy 955 GCTGTGTGCAGAACCCCGCTTCAGCCCGACCGCTGCGCTTATCCGGTAACTATCGTC 1014
Db 4308 gctgtgtgcagcaaccctccttcagcccgacctgcgcttaccggttaactatcgtc 4367
Qy 1015 TTGAGTCCAAACCCGGTAAGACACGACTTATCGCCACTGGCAGCAGCCACTGGTAAACAGGA 1074
Db 4368 ttgagtcacaaccggtaagacacgacttatcgccactggcagcagcactggttaacagga 4427
Qy 1075 TTAGCAGAGCGAGGTATGTAGCGGTGCTACAGAGTTCTTCAAGTGTGGCTAACTACG 1134
Db 4428 ttgacagagcgaggtatgtagcggtgtcagaggtcttgaaagtggcgctaactacg 4487
Qy 1135 GCTACACTAGAAAGACAGTATTTGGTATCTGCGTCTGCTGAAGCCAGTTACCTTCGGAA 1194
Db 4488 gctacactagaagacagatttggtagtctgctgctgctgaagcagttaccttcggaa 4547
Qy 1195 AAAGAGTGTGTAGTCTTGTATCCGGCAAAACAAACACCGCTGGTAGCGGTGGTTTTTGTG 1254
Db 4548 aaagagttggtagctcttgatcccgcaaaacaccccgctggtagcggtggtgtttttg 4607
Qy 1255 TTTCGAACGACAGATACGCGCGCAGAAAAAGGATCT 1292
Db 4608 ttgcaagcagcagattacgcgcgcaaaaaaaggatct 4645
```

[illegible]

```

RESULT      4
US-09-011-745-7
; Sequence 7, Application US/09011745
; Patent No. 6165715
; GENERAL INFORMATION:
; APPLICANT: Collins, Mary KL
; APPLICANT: Weiss, Robin A
; APPLICANT: Takeuchi, Yasuhiro
; APPLICANT: Cosset, Francois-Loic
; TITLE OF INVENTION: Expression systems
; FILE REFERENCE: 09/011,745
; CURRENT APPLICATION NUMBER: US/09/011,745
; EARLIER FILING DATE: 1998-06-22
; EARLIER APPLICATION NUMBER: PCT/GB96/02061
; EARLIER FILING DATE: 1996-08-23
; EARLIER APPLICATION NUMBER: GB9517263.1
; EARLIER FILING DATE: 1995-08-23
; NUMBER OF SEQ ID NOS: 29
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 7
; LENGTH: 6312
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: Portion of
; OTHER INFORMATION: Construct
; FEATURE:
; NAME/KEY: misc.feature
; LOCATION: (4058)
; OTHER INFORMATION: n is any nucleotide
; FEATURE:
; NAME/KEY: misc.feature
; LOCATION: (4059)
; OTHER INFORMATION: n is any nucleotide
; FEATURE:
; NAME/KEY: misc.feature
; LOCATION: (4060)
; OTHER INFORMATION: n is any nucleotide
; FEATURE:
; NAME/KEY: misc.feature
; LOCATION: (4061)
; OTHER INFORMATION: n is any nucleotide
; FEATURE:
; NAME/KEY: misc.feature
; LOCATION: (4246)
; OTHER INFORMATION: n is any nucleotide
; FEATURE:
; NAME/KEY: misc.feature
; LOCATION: (4247)
; OTHER INFORMATION: n is any nucleotide
; FEATURE:
; NAME/KEY: misc.feature
; LOCATION: (4248)
; OTHER INFORMATION: n is any nucleotide
; FEATURE:
; NAME/KEY: misc.feature
; LOCATION: (4249)
; OTHER INFORMATION: n is any nucleotide
;
US-09-011-745-7

```

```

; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/954,148
; FILING DATE: 30-SEP-1992
; ATTORNEY/AGENT INFORMATION:
; NAME: Fitting, Thomas
; REGISTRATION NUMBER: 34,163
; REFERENCE/DOCKET NUMBER: SCR1452P
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 619-554-2937
; TELEFAX: 619-554-6312
; INFORMATION FOR SEQ ID NO: 156:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 13254 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: circular
; MOLECULE TYPE: DNA (genomic)
; US-08-276-852-156

Query Match 33.4%; Score 604.2; DB 1; Length 13254;
Best Local Similarity 97.9%; Pred. No. 1.6e-174;
Matches 612; Conservative 0; Mismatches 13; Indels 0; Gaps 0;

QY 668 TGAGAAAGAAATTCCTGACGCCCGCGGTTGCTGGCGTTTTTCCATAGGCTCCGCCCC 727
Db 5976 TAAGGAGAGCGTCGACCTCGGGCGGTTGCTGGCGTTTTTCCATAGGCTCCGCCCC 6035

QY 728 TGACGAGCATCACAAAATCGAGCTCAAGTCAGAGTGGCGAACCAGGACGACTATA 787
Db 6036 TGACGAGCATCACAAAATCGAGCTCAAGTCAGAGTGGCGAACCAGGACGACTATA 6095

QY 788 AAGATACAGCGGTTCCCTCCCTGSAAGCTCCCTCGTCGCTCTCCTGTCGACCCGTC 847
Db 6096 AAGATACAGCGGTTCCCTCCCTGSAAGCTCCCTCGTCGCTCTCCTGTCGACCCGTC 6155

QY 848 GCTTACCGGATACCTGTCGCCCTTCTCCCTTCGGGAAGCGTGGCGTTTCTCAATGCTC 907
Db 6156 GCTTACCGGATACCTGTCGCCCTTCTCCCTTCGGGAAGCGTGGCGTTTCTCAATGCTC 6215

QY 908 ACGTGTAGTATCTCAGTTCGGTGTAGTTCGTTTCGCTCCAAAGTCGGCTGTGTGCACGA 967
Db 6216 ACGTGTAGTATCTCAGTTCGGTGTAGTTCGTTTCGCTCCAAAGTCGGCTGTGTGCACGA 6275

QY 968 ACCCCCGTTACGCGGACCGCTGCGCTTATCCGGTAACTATCGTTCGATCCCAACC 1027
Db 6276 ACCCCCGTTACGCGGACCGCTGCGCTTATCCGGTAACTATCGTTCGATCCCAACC 6335

QY 1028 GGTAAAGACAGACTTATCGCCACTGGCAGCAGCCACTGGTAAACAGGATTAGCAGAGCGAG 1087
Db 6336 GGTAAAGACAGACTTATCGCCACTGGCAGCAGCCACTGGTAAACAGGATTAGCAGAGCGAG 6395

QY 1088 GTATGTAGCGGTGTACAGAGTTCCTTGAAGTGGTGGCGCTTAACCTACGGCTACACTAGAAG 1147
Db 6396 GTATGTAGCGGTGTACAGAGTTCCTTGAAGTGGTGGCGCTTAACCTACGGCTACACTAGAAG 6455

QY 1148 GACAGTATTTGGTATCTGCGCTCTGCTGAAGCCAGTTACCTTCGGAAGAAAGATTGTAG 1207
Db 6456 GACAGTATTTGGTATCTGCGCTCTGCTGAAGCCAGTTACCTTCGGAAGAAAGATTGTAG 6515

QY 1208 CTCTTGATCGGCAAAACAAACACCGCTGCTAGCGGTGGTTTTTTTGTTCGAAGCAGCA 1267
Db 6516 CTCTTGATCGGCAAAACAAACACCGCTGCTAGCGGTGGTTTTTTTGTTCGAAGCAGCA 6575

QY 1268 GATTACCGCGAGAAAAAAGGATCT 1292
Db 6576 GATTACCGCGAGAAAAAAGGATCT 6600

```

RESULT 6
 US-08-276-852-170/c
 ; Sequence 170, Application US/08276852
 ; Patent No. 5652138

```

; GENERAL INFORMATION:
; APPLICANT: Burton, Dennis R
; APPLICANT: Bernas, Carlos F
; APPLICANT: Lerner, Richard A
; TITLE OF INVENTION: HUMAN NEUTRALIZING MONOCLONAL ANTIBODIES
; TITLE OF INVENTION: TO HUMAN IMMUNODEFICIENCY VIRUS
; NUMBER OF SEQUENCES: 170
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: The Scripps Research Institute, Office of
; ADDRESSEE: Patent Counsel
; STREET: 10666 No. 5652138th Torrey Pines Road, Suite 220,
; STREET: Mail Drop TPC8
; CITY: La Jolla
; STATE: CA
; COUNTRY: USA
; ZIP: 92037
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA: US/08/276,852
; FILING DATE: 18-JUL-1994
; CLASSIFICATION: 514
; PRIOR APPLICATION NUMBER: US 08/178,302
; FILING DATE: 30-SEP-1993
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/954,148
; FILING DATE: 30-SEP-1992
; ATTORNEY/AGENT INFORMATION:
; NAME: Fitting, Thomas
; REGISTRATION NUMBER: 34,163
; REFERENCE/DOCKET NUMBER: SCR1452P
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 619-554-2937
; TELEFAX: 619-554-6312
; INFORMATION FOR SEQ ID NO: 170:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 13254 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: circular
; MOLECULE TYPE: DNA (genomic)
; US-08-276-852-170

```

```

Query Match 33.4%; Score 604.2; DB 1; Length 13254;
Best Local Similarity 97.9%; Pred. No. 1.6e-174;
Matches 612; Conservative 0; Mismatches 13; Indels 0; Gaps 0;

QY 668 TGAGAAAGAAATTCCTGACGCCCGCGGTTGCTGGCGTTTTTCCATAGGCTCCGCCCC 727
Db 7279 TAAGGAGAGCGTCGACCTCGGGCGGTTGCTGGCGTTTTTCCATAGGCTCCGCCCC 7220

QY 728 TGACGAGCATCACAAAATCGAGCTCAAGTCAGAGTGGCGAACCAGGACGACTATA 787
Db 7219 TGACGAGCATCACAAAATCGAGCTCAAGTCAGAGTGGCGAACCAGGACGACTATA 7160

QY 788 AAGATACAGCGGTTTCCCTCCCTGGAAGCTCCCTCGTCGCTCTCCTGTCGACCCGTC 847
Db 7159 AAGATACAGCGGTTTCCCTCCCTGGAAGCTCCCTCGTCGCTCTCCTGTCGACCCGTC 7100

QY 848 GCTTACCGGATACCTGTCGCCCTTCTCCCTTCGGGAAGCGTGGCGTTTCTCAATGCTC 907
Db 7099 GCTTACCGGATACCTGTCGCCCTTCTCCCTTCGGGAAGCGTGGCGTTTCTCAATGCTC 7040

QY 908 ACGTGTAGTATCTCAGTTCGGTGTAGTTCGCTCCAAAGCTCGGCTGTGTGCACGA 967
Db 7039 ACGTGTAGTATCTCAGTTCGGTGTAGTTCGCTCCAAAGCTCGGCTGTGTGCACGA 6980

QY 968 ACCCCCGTTACGCGGACCGCTGCGCTTATCCGGTAACTATCGTTCGATCCCAACC 1027

```


; Sequence 156, Application PC/TUS9508743
; GENERAL INFORMATION:
; APPLICANT:
; TITLE OF INVENTION: HUMAN NEUTRALIZING MONOCLONAL ANTIBODIES
; TITLE OF INVENTION: TO HUMAN IMMUNODEFICIENCY VIRUS
; NUMBER OF SEQUENCES: 170
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25 (EPO)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: PCT/US95/08743
; FILING DATE: 11-JUL-1995
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/276,852
; FILING DATE: 18-JUL-1994
; INFORMATION FOR SEQ ID NO: 156:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 13254 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: circular
; MOLECULE TYPE: DNA (genomic)
; PCT-US95-08743-156

Query Match 33.4%; Score 604.2; DB 5; Length 13254;
Best Local Similarity 97.9%; Pred. No. 1.6e-174;
Matches 612; Conservative 0; Mismatches 13; Indels 0; Gaps 0;

QY 668 TGAGAAAGATTCCTGCAGCGCGCGTGGCGCTTTTCATAGGCTCGCGCCCC 727
DB 5976 TAAGGAGAGCGTCGACCTCGGCGCGGTGGCGCTTTTCATAGGCTCGCGCCCC 6035

QY 728 TGACGAGCATCACAAAATCGAGCTCAAGTCAAGTGGCGAAGCCGACGACTATA 787
DB 6036 TGACGAGCATCACAAAATCGAGCTCAAGTCAAGTGGCGAAGCCGACGACTATA 6095

QY 788 AAGATACCGAGCGTTTCCCGCTGGAAGCTCCCTCGCGCTCTCTGTTCCGACCTGCC 847
DB 6096 AAGATACCGAGCGTTTCCCGCTGGAAGCTCCCTCGCGCTCTCTGTTCCGACCTGCC 6155

QY 848 GCTTACCGGATACCTGTCCGCTTTCCTCCCTCGGGAAGCGTGGCGCTTCTCAATGCTC 907
DB 6156 GCTTACCGGATACCTGTCCGCTTTCCTCCCTCGGGAAGCGTGGCGCTTCTCAATGCTC 6215

QY 908 ACGCTGTAGTATCTCAGTTCGGTGTAGTTCGCTTCCAAAGCTGGGCTGTGTGCACGA 967
DB 6216 ACGCTGTAGTATCTCAGTTCGGTGTAGTTCGCTTCCAAAGCTGGGCTGTGTGCACGA 6275

QY 968 ACCCGCGTTACGCGCGCGCTGGCGCTTATCCGGTAACTATCGTCTTGAGTCCACACC 1027
DB 6276 ACCCGCGTTACGCGCGCGCTGGCGCTTATCCGGTAACTATCGTCTTGAGTCCACACC 6335

QY 1028 GGTAAAGACAGCACTTATCGCACTGGCAGAGCCACCTGGTAAAGAGATTAGCAGAGCGAG 1087
DB 6336 GGTAAAGACAGCACTTATCGCACTGGCAGAGCCACCTGGTAAAGAGATTAGCAGAGCGAG 6395

QY 1088 GTATGTAGGCGGTGCTACAGAGTTCCTTGAAGTGGTGGCGCTTAACCTACGCTACACTAGAAG 1147
DB 6396 GTATGTAGGCGGTGCTACAGAGTTCCTTGAAGTGGTGGCGCTTAACCTACGCTACACTAGAAG 6455

QY 1148 GACAGTATTGGTATCTCGCGCTGTGTAAGCCAGTTACCTTCGGAAAAAGAGTTGGTAG 1207
DB 6456 GACAGTATTGGTATCTCGCGCTGTGTAAGCCAGTTACCTTCGGAAAAAGAGTTGGTAG 6515

QY 1208 CTCCTGTATCGGCAACAAACACCGCTGCTAGCGGTGGTGTGTTTGTGTCAGACGACGA 1267
DB 6516 CTCCTGTATCGGCAACAAACACCGCTGCTAGCGGTGGTGTGTTTGTGTCAGACGACGA 6575

QY 1268 GATTACCGCGCAAAAAAAGGATCT 1292
|||||

Db 6576 GATTACCGCGCAAAAAAAGGATCT 6600

RESULT 12
PCT-US95-08743-170/c
; Sequence 170, Application PC/TUS9508743
; GENERAL INFORMATION:
; APPLICANT:
; TITLE OF INVENTION: HUMAN NEUTRALIZING MONOCLONAL ANTIBODIES
; TITLE OF INVENTION: TO HUMAN IMMUNODEFICIENCY VIRUS
; NUMBER OF SEQUENCES: 170
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25 (EPO)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: PCT/US95/08743
; FILING DATE: 11-JUL-1995
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/276,852
; FILING DATE: 18-JUL-1994
; INFORMATION FOR SEQ ID NO: 170:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 13254 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: circular
; MOLECULE TYPE: DNA (genomic)
; PCT-US95-08743-170

Query Match 33.4%; Score 604.2; DB 5; Length 13254;
Best Local Similarity 97.9%; Pred. No. 1.6e-174;
Matches 612; Conservative 0; Mismatches 13; Indels 0; Gaps 0;

QY 668 TGAGAAAGATTCCTGCAGCGCGCGTGGCGCTTTTCATAGGCTCGCGCCCC 727
DB 7279 TAAGGAGAGCGTCGACCTCGGCGCGGTGGCGCTTTTCATAGGCTCGCGCCCC 7220

QY 728 TGACGAGCATCACAAAATCGAGCTCAAGTCAAGTGGCGAAGCCGACGACTATA 787
DB 7219 TGACGAGCATCACAAAATCGAGCTCAAGTCAAGTGGCGAAGCCGACGACTATA 7160

QY 788 AAGATACCGAGCGTTTCCCGCTGGAAGCTCCCTCGCGCTCTCTGTTCCGACCTGCC 847
DB 7159 AAGATACCGAGCGTTTCCCGCTGGAAGCTCCCTCGCGCTCTCTGTTCCGACCTGCC 7100

QY 848 GCTTACCGGATACCTGTCCGCTTTCCTCCCTCGGGAAGCGTGGCGCTTCTCAATGCTC 907
DB 7099 GCTTACCGGATACCTGTCCGCTTTCCTCCCTCGGGAAGCGTGGCGCTTCTCAATGCTC 7040

QY 908 ACGCTGTAGTATCTCAGTTCGGTGTAGTTCGCTTCCAAAGCTGGGCTGTGTGCACGA 967
DB 7039 ACGCTGTAGTATCTCAGTTCGGTGTAGTTCGCTTCCAAAGCTGGGCTGTGTGCACGA 6980

QY 968 ACCCGCGTTACGCGCGCGCTGGCGCTTATCCGGTAACTATCGTCTTGAGTCCACACC 1027
DB 6979 ACCCGCGTTACGCGCGCGCTGGCGCTTATCCGGTAACTATCGTCTTGAGTCCACACC 6920

QY 1028 GGTAAAGACAGCACTTATCGCACTGGCAGAGCCACCTGGTAAAGAGATTAGCAGAGCGAG 1087
DB 6919 GGTAAAGACAGCACTTATCGCACTGGCAGAGCCACCTGGTAAAGAGATTAGCAGAGCGAG 6860

QY 1088 GTATGTAGGCGGTGCTACAGAGTTCCTTGAAGTGGTGGCGCTTAACCTACGCTACACTAGAAG 1147
DB 6859 GTATGTAGGCGGTGCTACAGAGTTCCTTGAAGTGGTGGCGCTTAACCTACGCTACACTAGAAG 6800

QY 1148 GACAGTATTGGTATCTCGCGCTGTGTAAGCCAGTTACCTTCGGAAAAAGAGTTGGTAG 1207
DB 6799 GACAGTATTGGTATCTCGCGCTGTGTAAGCCAGTTACCTTCGGAAAAAGAGTTGGTAG 6740

QY 1208 CTCCTGTATCGGCAACAAACACCGCTGCTAGCGGTGGTGTGTTTGTGTCAGACGACGA 1267
|||||

Db 247 TTTCTCCCTCGGGAAGCGTGGCGCTTTCTCAATGCTCAGCTGTAGGTATCTCAGTTGG 306
QY 930 GTGTAGTGTGCTCGCTCCCAAGCTGGGCTGTGTGACGAAACCCCGCTTCAGCCCGACCGC 989
Db 307 GTGTAGTGTGCTCGCTCCCAAGCTGGGCTGTGTGACGAAACCCCGCTTCAGCCCGACCGC 366
QY 990 TGGCGCTTATCCGTAAGTATCTGTGTTGAGTCCAAACCCCGTAAGACACACTTATCGCCA 1049
Db 367 TGGCGCTTATCCGTAAGTATCTGTGTTGAGTCCAAACCCCGTAAGACACACTTATCGCCA 426
QY 1050 CTGCGACGACGCTGTTGTAACAGATTAGCAGCGGAGGTATGTAGCGGTGTACAGAG 1109
Db 427 CTGCGACGACGCTGTTGTAACAGATTAGCAGCGGAGGTATGTAGCGGTGTACAGAG 486
QY 1110 TTCTTGAAGTGTGGCTTAACCTACGCTACACTAGAGGACAGTATTTGGTATCTCGGCT 1169
Db 487 TTCTTGAAGTGTGGCTTAACCTACGCTACACTAGAGGACAGTATTTGGTATCTCGGCT 546
QY 1170 CTGCTGAAGCCAGTTACTTTCGGAAGAAAGAGTTGGTAGCTCTTCATCCGCAACAAACC 1229
Db 547 CTGCTGAAGCCAGTTACTTTCGGAAGAAAGAGTTGGTAGCTCTTCATCCGCAACAAACC 606
QY 1230 ACCGCTGTAGCGGTGTTTTTTTGTGCAAGCAGCAGATTTAGCGCAGAAAAAGGA 1289
Db 607 ACCGCTGTAGCGGTGTTTTTTTGTGCAAGCAGCAGATTTAGCGCAGAAAAAGGA 666
QY 1290 TCT 1292
Db 667 TCT 669
RESULT 14
US-08-906-957-9
; Sequence 9, Application US/08906957
; Patent No. 5856142
; GENERAL INFORMATION:
; APPLICANT: LEGOUX, Richard
; APPLICANT: MALDONADO, Paul
; TITLE OF INVENTION: Method for the extraction of
; TITLE OF INVENTION: periplasmic proteins of prokaryotic microorganisms in the
; TITLE OF INVENTION: presence of arginine
; NUMBER OF SEQUENCES: 9
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Bacon & Thomas
; STREET: 625 Slaters Lane - Fourth Floor
; CITY: Alexandria
; STATE: Virginia
; COUNTRY: USA
; ZIP: 22314
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/906,957
; FILING DATE: 06-AUG-1997
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/594,469
; FILING DATE:
; APPLICATION NUMBER: FR 95 01083
; FILING DATE: 31-JAN-1995
; ATTORNEY/AGENT INFORMATION:
; NAME: FICHTER, Richard E
; REGISTRATION NUMBER: 26,382
; REFERENCE/DOCKET NUMBER: REF/LEGOUX
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (703) 683-0500
; TELEFAX: (703) 683-1080
; INFORMATION FOR SEQ ID NO: 9:

Db 6739 CTCCTGATCCGCAACAAACACCGCTGGTAGCGTGTGTTTTTTTGTGTCGAAGCAGCA 6680
QY 1268 GATTAGCGCGCAAAAAAAGGATCT 1292
Db 6679 GATTAGCGCGCAAAAAAAGGATCT 6655
RESULT 13
US-08-594-469-9
; Sequence 9, Application US/08594469
; Patent No. 5700665
; GENERAL INFORMATION:
; APPLICANT: LEGOUX, Richard
; APPLICANT: MALDONADO, Paul
; APPLICANT: SALOME, Marc
; TITLE OF INVENTION: Method for the extraction of
; TITLE OF INVENTION: periplasmic proteins of prokaryotic microorganisms in the
; TITLE OF INVENTION: presence of arginine
; NUMBER OF SEQUENCES: 9
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Bacon & Thomas
; STREET: 625 Slaters Lane - Fourth Floor
; CITY: Alexandria
; STATE: Virginia
; COUNTRY: USA
; ZIP: 22314
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/594,469
; FILING DATE:
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: FR 95 01083
; FILING DATE: 31-JAN-1995
; ATTORNEY/AGENT INFORMATION:
; NAME: FICHTER, Richard E
; REGISTRATION NUMBER: 26,382
; REFERENCE/DOCKET NUMBER: REF/LEGOUX
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (703) 683-0500
; TELEFAX: (703) 683-1080
; INFORMATION FOR SEQ ID NO: 9:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1905 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
US-08-594-469-9
Query Match 33.4%; Score 603; DB 1; Length 1905;
Best Local Similarity 100.0%; Pred. No. 1.3e-174;
Matches 603; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 690 GCGCGCTGCTGGCGTTTTCATAGGTCGCCGCCCTGAGGACATCAACAAATCGA 749
Db 67 GCGCGCTGCTGGCGTTTTCATAGGTCGCCGCCCTGAGGACATCAACAAATCGA 126
QY 750 CGTCAAGTCAGAGTGGCGAACCAGCAGACTATAAGATACAGCGTTTCCCGCT 809
Db 127 CGTCAAGTCAGAGTGGCGAACCAGCAGACTATAAGATACAGCGTTTCCCGCT 186
QY 810 GGAAGTCCCTCGTGGCGTCTCTGTTCCGACCGCTCCCGCTTACCGGATACCTGCCCGC 869
Db 187 GGAAGTCCCTCGTGGCGTCTCTGTTCCGACCGCTCCCGCTTACCGGATACCTGCCCGC 246
QY 870 TTCTCCCTTCGGGAAGCGTGGCGCTTTCTCAATGCTCAGCTGTAGGTATCTCAGTTGG 929

; SEQUENCE CHARACTERISTICS:
; LENGTH: 1905 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
US-08-906-957-9

Query Match 33.4%; Score 603; DB 2; Length 1905;
Best Local Similarity 100.0%; Pred. No. 1.3e-174;
Matches 603; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 690 GCCCGGTTGCTGGCGTTTTCATAGCTCGCCCGCTGACGAGCATCAAAAAATCGA 749
DB 67 GCCCGGTTGCTGGCGTTTTCATAGCTCGCCCGCTGACGAGCATCAAAAAATCGA 126

QY 750 CGCTCAAGTCAGAGGTGGGAAACCCGACAGGACTATAAAGATACCAAGGCGTTTCCCGCT 809
DB 127 CGCTCAAGTCAGAGGTGGGAAACCCGACAGGACTATAAAGATACCAAGGCGTTTCCCGCT 186

QY 810 GGAAGCTCCCTCGTGGCGTCTCTGTTCCGACCTGCGGCTTACCGGATACCTGTCGCC 869
DB 187 GGAAGCTCCCTCGTGGCGTCTCTGTTCCGACCTGCGGCTTACCGGATACCTGTCGCC 246

QY 870 TTTCTCCCTCGGGAAGCGTGGCGTTTCTCAATGCTCAGCTGTAGGTATCTCAGTTGCG 929
DB 247 TTTCTCCCTCGGGAAGCGTGGCGTTTCTCAATGCTCAGCTGTAGGTATCTCAGTTGCG 306

QY 930 GTGAGTGTGCTCGCTCCAAAGCTGGGCTGTGTGCAGAACCCCGCTTCAGCCGACCGC 989
DB 307 GTGAGTGTGCTCGCTCCAAAGCTGGGCTGTGTGCAGAACCCCGCTTCAGCCGACCGC 366

QY 990 TGCCTTATCCGCTAATATCTGTTGAGTCCAAACCCGGTAAAGACGAGCTTATCGCCA 1049
DB 367 TGCCTTATCCGCTAATATCTGTTGAGTCCAAACCCGGTAAAGACGAGCTTATCGCCA 426

QY 1050 CTGCAGCAGCCACTGTTAAGAGATTAGCAGAGGATGTAGCGGTGTACAGAG 1109
DB 427 CTGCAGCAGCCACTGTTAAGAGATTAGCAGAGGATGTAGCGGTGTACAGAG 486

QY 1110 TTTCTGAAGTGGTGGCTTAAGTACCGGTACACTAGAAGGACAGATTGTTGATCTGCGCT 1169
DB 487 TTTCTGAAGTGGTGGCTTAAGTACCGGTACACTAGAAGGACAGATTGTTGATCTGCGCT 546

QY 1170 CTGCTGAAGCCAGTTACCTTCGGAAGAGATTGGTAGCTCTTATCGGCAAAAC 1229
DB 547 CTGCTGAAGCCAGTTACCTTCGGAAGAGATTGGTAGCTCTTATCGGCAAAAC 606

QY 1230 ACCGCTGTAGCGTGGTGTGTTTGTGCAAGCAGAGATTACCGCGCAGAAAAAGGA 1289
DB 607 ACCGCTGTAGCGTGGTGTGTTTGTGCAAGCAGAGATTACCGCGCAGAAAAAGGA 666

QY 1290 TCT 1292
DB 667 TCT 669

RESULT. 15
US-08-941-647A-1/c
; Sequence 1, Application us/08941647A
; Patent No. 5952211
; GENERAL INFORMATION:
; APPLICANT: NAKAMURA, ATSUGO
; APPLICANT: TANAKA, TOSHIO
; APPLICANT: MATSUO, YUSHI
; APPLICANT: TANASE, SUMIO
; APPLICANT: FUNATSU, MASAHICO
; APPLICANT: ETO, AKIRA
; TITLE OF INVENTION: METHOD FOR PRODUCING ACTIVE HUMAN
; NUMBER OF SEQUENCES: 4
; CORRESPONDENCE ADDRESS:

; ADDRESSEE: OBLON, SPIVAK, MCCLELLAND, MAIER & NEUSTADT,
; ADDRESSEE: P.C.
; STREET: 1755 S. JEFFERSON DAVIS HIGHWAY, SUITE 400
; CITY: ARLINGTON
; STATE: VA
; COUNTRY: USA
; ZIP: 22202
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/941,647A
; FILING DATE: 30-SEP-1997
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/533,717
; FILING DATE: 26-SEP-1995
; APPLICATION NUMBER: JP 6-268119
; FILING DATE: 07-OCT-1994
; ATTORNEY/AGENT INFORMATION:
; NAME: OBLON, NORMAN F.
; REGISTRATION NUMBER: 24,618
; REFERENCE/DOCKET NUMBER: 7335-001-0
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 703-413-3000
; TELEFAX: 703-413-2220
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2927 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
US-08-941-647A-1

Query Match 33.4%; Score 603; DB 2; Length 2927;
Best Local Similarity 100.0%; Pred. No. 1.6e-174;
Matches 603; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 690 GCCCGGTTGCTGGCGTTTTCATAGCTCGCCCGCTGACGAGCATCAAAAAATCGA 749
DB 2763 GCCCGGTTGCTGGCGTTTTCATAGCTCGCCCGCTGACGAGCATCAAAAAATCGA 2704

QY 750 CGCTCAAGTCAGAGGTGGGAAACCCGACAGGACTATAAAGATACCAAGGCGTTTCCCGCT 809
DB 2703 CGCTCAAGTCAGAGGTGGGAAACCCGACAGGACTATAAAGATACCAAGGCGTTTCCCGCT 2644

QY 810 GGAAGCTCCCTCGTGGCGTCTCTGTTCCGACCTGCGGCTTACCGGATACCTGTCGCC 869
DB 2643 GGAAGCTCCCTCGTGGCGTCTCTGTTCCGACCTGCGGCTTACCGGATACCTGTCGCC 2584

QY 870 TTTCTCCCTCGGGAAGCGTGGCGTTTCTCAATGCTCACGCTGTAGGTATCTCAGTTGCG 929
DB 2583 TTTCTCCCTCGGGAAGCGTGGCGTTTCTCAATGCTCACGCTGTAGGTATCTCAGTTGCG 2524

QY 930 GTGAGTGTGCTCGCTCCAAAGCTGGGCTGTGTGCAGAACCCCGCTTCAGCCGACCGC 989
DB 2523 GTGAGTGTGCTCGCTCCAAAGCTGGGCTGTGTGCAGAACCCCGCTTCAGCCGACCGC 2464

QY 990 TGCCTTATCCGCTAATATCTGTTGAGTCCAAACCCGGTAAAGACGAGCTTATCGCCA 1049
DB 2463 TGCCTTATCCGCTAATATCTGTTGAGTCCAAACCCGGTAAAGACGAGCTTATCGCCA 2404

QY 1050 CTGCAGCAGCCACTGTTAAGAGATTAGCAGAGGATGTAGCGGTGTACAGAG 1109
DB 2403 CTGCAGCAGCCACTGTTAAGAGATTAGCAGAGGATGTAGCGGTGTACAGAG 2344

QY 1110 TTTCTGAAGTGGTGGCTTAAGTACCGGTACACTAGAAGGACAGATTGTTGATCTGCGCT 1169
DB 2343 TTTCTGAAGTGGTGGCTTAAGTACCGGTACACTAGAAGGACAGATTGTTGATCTGCGCT 2284

QY 1170 CTGCTGAAGCAGTTACCTTCGGAAGAGTTGGTAGCTCTTGATCCGGCAACAAACC 1229
|||||
Db 2283 CTGCTGAAGCCAGTTACCTTCGGAAGAGTTGGTAGCTCTTGATCCGGCAACAAACC 2224
|||||
QY 1230 ACCGCTGGTAGCGGTGTTTTTTTTCACAGCAGCAGATTACGGCGCAGAAAAAGGA 1289
|||||
Db 2223 ACCGCTGGTAGCGGTGTTTTTTTTCACAGCAGCAGATTACGGCGCAGAAAAAGGA 2164
|||||
QY 1290 TCT 1292
|||
Db 2163 TCT 2161

Search completed: January 17, 2002, 11:57:05
Job time: 17552 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: January 17, 2002, 15:14:01 ; Search time 11375.6 Seconds
(without alignments)
2784.055 Million cell updates/sec

Title: US-09-242-202A-28
Perfect score: 1807
Sequence: 1 GGTACTGTCACCATGGCGC.....CTCGACAGCCTCTCCACA 1807

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0
Searched: 17159718 seqs, 8763200856 residues
Total number of hits satisfying chosen parameters: 34319436

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Pending_Patents_NA_Main:*

1:	/cgn2_6/ptodata/2/pna/PCTUS_COMB.seq:*
2:	/cgn2_6/ptodata/2/pna/US06_COMB.seq:*
3:	/cgn2_6/ptodata/2/pna/US07_COMB.seq:*
4:	/cgn2_6/ptodata/2/pna/US08_COMB.seq:*
5:	/cgn2_6/ptodata/2/pna/US081_COMB.seq:*
6:	/cgn2_6/ptodata/2/pna/US082_COMB.seq:*
7:	/cgn2_6/ptodata/2/pna/US083_COMB.seq:*
8:	/cgn2_6/ptodata/2/pna/US084_COMB.seq:*
9:	/cgn2_6/ptodata/2/pna/US085_COMB.seq:*
10:	/cgn2_6/ptodata/2/pna/US086_COMB.seq:*
11:	/cgn2_6/ptodata/2/pna/US087_COMB.seq:*
12:	/cgn2_6/ptodata/2/pna/US088_COMB.seq:*
13:	/cgn2_6/ptodata/2/pna/US089_COMB.seq:*
14:	/cgn2_6/ptodata/2/pna/US090_COMB.seq:*
15:	/cgn2_6/ptodata/2/pna/US091_COMB.seq:*
16:	/cgn2_6/ptodata/2/pna/US092_COMB.seq:*
17:	/cgn2_6/ptodata/2/pna/US093_COMB.seq:*
18:	/cgn2_6/ptodata/2/pna/US094_COMB.seq:*
19:	/cgn2_6/ptodata/2/pna/US095A_COMB.seq:*
20:	/cgn2_6/ptodata/2/pna/US095B_COMB.seq:*
21:	/cgn2_6/ptodata/2/pna/US095C_COMB.seq:*
22:	/cgn2_6/ptodata/2/pna/US095D_COMB.seq:*
23:	/cgn2_6/ptodata/2/pna/US096A_COMB.seq:*
24:	/cgn2_6/ptodata/2/pna/US096B_COMB.seq:*
25:	/cgn2_6/ptodata/2/pna/US096C_COMB.seq:*
26:	/cgn2_6/ptodata/2/pna/US096D_COMB.seq:*
27:	/cgn2_6/ptodata/2/pna/US096E_COMB.seq:*
28:	/cgn2_6/ptodata/2/pna/US097A_COMB.seq:*
29:	/cgn2_6/ptodata/2/pna/US097B_COMB.seq:*
30:	/cgn2_6/ptodata/2/pna/US097C_COMB.seq:*
31:	/cgn2_6/ptodata/2/pna/US098_COMB.seq:*
32:	/cgn2_6/ptodata/2/pna/US099_COMB.seq:*
33:	/cgn2_6/ptodata/2/pna/US6000_COMB.seq:*
34:	/cgn2_6/ptodata/2/pna/US6001_COMB.seq:*
35:	/cgn2_6/ptodata/2/pna/US6002_COMB.seq:*
36:	/cgn2_6/ptodata/2/pna/US6003_COMB.seq:*
37:	/cgn2_6/ptodata/2/pna/US6004_COMB.seq:*
38:	/cgn2_6/ptodata/2/pna/US6005_COMB.seq:*
39:	/cgn2_6/ptodata/2/pna/US6006_COMB.seq:*
40:	/cgn2_6/ptodata/2/pna/US6007_COMB.seq:*
41:	/cgn2_6/ptodata/2/pna/US6008_COMB.seq:*
42:	/cgn2_6/ptodata/2/pna/US6009_COMB.seq:*
43:	/cgn2_6/ptodata/2/pna/US6010_COMB.seq:*

44:	/cgn2_6/ptodata/2/pna/US6011_COMB.seq:*
45:	/cgn2_6/ptodata/2/pna/US6012_COMB.seq:*
46:	/cgn2_6/ptodata/2/pna/US6013_COMB.seq:*
47:	/cgn2_6/ptodata/2/pna/US6014_COMB.seq:*
48:	/cgn2_6/ptodata/2/pna/US6015_COMB.seq:*
49:	/cgn2_6/ptodata/2/pna/US6016_COMB.seq:*
50:	/cgn2_6/ptodata/2/pna/US6017_COMB.seq:*
51:	/cgn2_6/ptodata/2/pna/US6018_COMB.seq:*
52:	/cgn2_6/ptodata/2/pna/US6019_COMB.seq:*
53:	/cgn2_6/ptodata/2/pna/US6020_COMB.seq:*
54:	/cgn2_6/ptodata/2/pna/US6021_COMB.seq:*
55:	/cgn2_6/ptodata/2/pna/US6022_COMB.seq:*
56:	/cgn2_6/ptodata/2/pna/US6023_COMB.seq:*
57:	/cgn2_6/ptodata/2/pna/US6024_COMB.seq:*
58:	/cgn2_6/ptodata/2/pna/US6025_COMB.seq:*
59:	/cgn2_6/ptodata/2/pna/US6026_COMB.seq:*
60:	/cgn2_6/ptodata/2/pna/US6027_COMB.seq:*
61:	/cgn2_6/ptodata/2/pna/US6028_COMB.seq:*
62:	/cgn2_6/ptodata/2/pna/US6029_COMB.seq:*
63:	/cgn2_6/ptodata/2/pna/US6030_COMB.seq:*
64:	/cgn2_6/ptodata/2/pna/US6031_COMB.seq:*
65:	/cgn2_6/ptodata/2/pna/US6032_COMB.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB	ID	Description
1	1807	100.0	1807	16	US-09-242-202-28	Sequence 28, Appl
2	1807	100.0	1807	16	US-09-242-202A-28	Sequence 28, Appl
3	1583.2	87.6	2308	16	US-09-242-202-29	Sequence 29, Appl
4	1583.2	87.6	2308	16	US-09-242-202A-29	Sequence 29, Appl
5	1310.4	72.5	1547	16	US-09-242-202-27	Sequence 27, Appl
6	1310.4	72.5	1547	16	US-09-242-202A-27	Sequence 27, Appl
7	607.4	33.6	7252	16	US-09-238-356-27	Sequence 27, Appl
8	604.2	33.4	7073	8	US-08-480-120-15	Sequence 15, Appl
9	604.2	33.4	7864	8	US-08-480-120-20	Sequence 20, Appl
10	603	33.4	608	16	US-09-242-202-26	Sequence 26, Appl
11	603	33.4	608	16	US-09-242-202A-26	Sequence 26, Appl
12	603	33.4	2077	1	PCT-US99-06742-7	Sequence 7, Appl
13	603	33.4	2462	18	US-09-496-445-5	Sequence 5, Appl
14	603	33.4	2462	44	US-60-118-478-5	Sequence 5, Appl
15	603	33.4	2577	3	US-07-919-535C-26	Sequence 26, Appl
16	603	33.4	2704	24	US-09-623-551-27	Sequence 27, Appl
17	603	33.4	2754	5	US-08-114-691-7	Sequence 7, Appl
18	603	33.4	2927	9	US-08-533-717-1	Sequence 1, Appl
19	603	33.4	2927	15	US-09-142-481-14	Sequence 14, Appl
20	603	33.4	2939	1	PCT-US94-10146-2	Sequence 2, Appl
21	603	33.4	2939	8	US-08-488-015-2	Sequence 2, Appl
22	603	33.4	2939	15	US-09-114-835-2	Sequence 2, Appl
23	603	33.4	3175	1	PCT-US99-06704-2	Sequence 2, Appl
24	603	33.4	3175	25	US-09-646-075-2	Sequence 2, Appl
25	603	33.4	3301	8	US-08-447-430-42	Sequence 42, Appl
26	603	33.4	3301	17	US-09-342-673-42	Sequence 42, Appl
27	603	33.4	3327	21	US-09-546-411A-56	Sequence 56, Appl
28	603	33.4	3327	21	US-09-547-939A-56	Sequence 56, Appl
29	603	33.4	3327	24	US-09-628-730-56	Sequence 56, Appl
30	603	33.4	3327	45	US-60-128-830-56	Sequence 56, Appl
31	603	33.4	3331	15	US-09-148-920-1	Sequence 1, Appl
32	603	33.4	3331	18	US-09-422-569-8	Sequence 8, Appl
33	603	33.4	3344	18	US-09-546-411A-48	Sequence 48, Appl
34	603	33.4	3369	21	US-09-547-939A-48	Sequence 48, Appl
35	603	33.4	3369	21	US-09-628-730-48	Sequence 48, Appl
36	603	33.4	3369	45	US-60-128-830-48	Sequence 48, Appl
37	603	33.4	3421	15	US-09-643-192-12	Sequence 12, Appl
38	603	33.4	3421	25	US-09-130-115-31	Sequence 31, Appl
39	603	33.4	3423	8	US-08-447-430-40	Sequence 40, Appl
40	603	33.4	3423	17	US-09-342-673-40	Sequence 40, Appl
41	603	33.4	3423	17	US-09-342-673-40	Sequence 40, Appl

c 42 603 33.4 3423 18 US-09-485-737A-86 Sequence 86, Appl
c 43 603 33.4 3423 18 US-09-485-737B-86 Sequence 86, Appl
44 603 33.4 3440 3 US-07-671-058-7 Sequence 7, Appl
45 603 33.4 3462 29 US-09-742-373-10 Sequence 10, Appl

ALIGNMENTS

RESULT 1
US-09-242-202-28
; Sequence 28, Application US/09242202
; GENERAL INFORMATION:
; APPLICANT: Nelson, Edward L.
; Nelson, Peter J.
; TITLE OF INVENTION: NOVEL VECTOR FOR
; POLYNUCLEOTIDE VACCINES
; NUMBER OF SEQUENCES: 29
; CORRESPONDENCE ADDRESS:
; ADDRESSER: MORGAN & FINNEGAN, L.L.P.
; STREET: 345 PARK AVENUE
; CITY: NEW YORK
; STATE: NEW YORK
; COUNTRY: USA
; ZIP: 10154
; COMPUTER READABLE FORM:
; MEDIUM TYPE: DISKETTE, 3.5 INCH, 1.44 MB
; COMPUTER: IBM PC COMPATIBLE
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: MS WORD 97
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/242,202
; FILING DATE: 01-Nov-1999
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US60/023931
; FILING DATE: 14-AUG-1996
; ATTORNEY/AGENT INFORMATION:
; NAME: KATHRYN M. BROWN
; REGISTRATION NUMBER: 34556
; REFERENCE/DOCKET NUMBER: 2026-4236051
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 758-4800
; TELEFAX: (212) 751-6849
; TELEX: 421792
; INFORMATION FOR SEQ ID NO: 28:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1807 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: unknown
; MOLECULE TYPE: cdna
; HYPOTHETICAL: No
; ANTI-SENSE: No
; SEQUENCE DESCRIPTION: SEQ ID NO: 28:
US-09-242-202-28

Query Match 100.0%; Score 1807; DB 16; Length 1807;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1807; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GGTACCTGCCACCAAGCGCGGATCTTTATCACTGATAAGTTGGTGACATATTATGTT 60
Db 1 GGTACCTGCCACCAAGCGCGGATCTTTATCACTGATAAGTTGGTGACATATTATGTT 60
Qy 61 TATCAGTATAAAGTGTCACAGTACGACCAAGTTGCGCGGAATACAGTATCGGTGCGCG 120
Db 61 TATCAGTATAAAGTGTCACAGTACGACCAAGTTGCGCGGAATACAGTATCGGTGCGCG 120
Qy 121 CCCTGGAGTGTGAACGAGGTGCGCGTAGAGGGTCTGACGACACGCAAACTGGCGGAACG 180
Db 121 CCCTGGAGTGTGAACGAGGTGCGCGTAGAGGGTCTGACGACACGCAAACTGGCGGAACG 180

Qy 181 GTTGGGGTGTCAGCAGCGCGGCTTTACTGSCACTTCAGGAACAAGCGCGGCTTTAAGG 240
Db 181 GTTGGGGTGTCAGCAGCGCGGCTTTACTGSCACTTCAGGAACAAGCGCGGCTTTAAGG 240
Qy 241 GCCATATGGTGAGTGGATGCTTCAACCCAGCGGGGATGGGGAGACCTCTAGTCAGAG 300
Db 241 GCCATATGGTGAGTGGATGCTTCAACCCAGCGGGGATGGGGAGACCTCTAGTCAGAG 300
Qy 301 CCCCAGGACACAGCGCAATGCCGCTTCCTTCCCTGCGAGATGAGTGTAGTGCTC 360
Db 301 CCCCAGGACACAGCGCAATGCCGCTTCCTTCCCTGCGAGATGAGTGTAGTGCTC 360
Qy 361 TCCTGGCCCTGGAAGTTGCCACTCCAGTGCACACAGCCTTGTCTTAATAAAATAGTT 420
Db 361 TCCTGGCCCTGGAAGTTGCCACTCCAGTGCACACAGCCTTGTCTTAATAAAATAGTT 420
Qy 421 GCATCATTTTCTGCTGACTAGTGTCTCTAATAATATATAAGCTTGATATCGAATCTTT 480
Db 421 GCATCATTTTCTGCTGACTAGTGTCTCTAATAATATATAAGCTTGATATCGAATCTTT 480
Qy 481 CGGACTTTTGAAGTGTGGTGGGGAAGGATTCGAACCTTCGAAGTCGATGACGGC 540
Db 481 CGGACTTTTGAAGTGTGGTGGGGAAGGATTCGAACCTTCGAAGTCGATGACGGC 540
Qy 541 AGATTAGAGTCTGCTCCCTTTGGCCGCTCGGGAACCCACACGGGTAACTTTTACT 600
Db 541 AGATTAGAGTCTGCTCCCTTTGGCCGCTCGGGAACCCACACGGGTAACTTTTACT 600
Qy 601 GCGCTGCTCCCTTATCGGGAAGCGGGCGCATCATATCAATGACGCGCGCTGTAAAGT 660
Db 601 GCGCTGCTCCCTTATCGGGAAGCGGGCGCATCATATCAATGACGCGCGCTGTAAAGT 660
Qy 661 GTTACGTTGAGAAAGATTCCTGACGCGCGCGCTGCTGCGGCTTTTCATAGGCTCC 720
Db 661 GTTACGTTGAGAAAGATTCCTGACGCGCGCGCTGCTGCGGCTTTTCATAGGCTCC 720
Qy 721 GCCCCCTGACGAGCATCACAAAATCGACGCTCAAGTCAGAGTGGCGAACCACGACAG 780
Db 721 GCCCCCTGACGAGCATCACAAAATCGACGCTCAAGTCAGAGTGGCGAACCACGACAG 780
Qy 781 GACTATAAAGATACCAAGCGGTTTCCCTGGAAGTCCTCGTGGCTCTCTCTGTTCCGA 840
Db 781 GACTATAAAGATACCAAGCGGTTTCCCTGGAAGTCCTCGTGGCTCTCTCTGTTCCGA 840
Qy 841 CCCTGCCGCTTACCGGATACCTGCGGCTTTCTCCCTTCGGGAAGCGTGGCGCTTCTC 900
Db 841 CCCTGCCGCTTACCGGATACCTGCGGCTTTCTCCCTTCGGGAAGCGTGGCGCTTCTC 900
Qy 901 AATGCTCACGCTGTAGGTATCTCAGTTCGCTGAGTGTAGTTCGCTCCAAGCTGGGCTGTG 960
Db 901 AATGCTCACGCTGTAGGTATCTCAGTTCGCTGAGTGTAGTTCGCTCCAAGCTGGGCTGTG 960
Qy 961 TGACGAAACCCCGCTTACGCGCGGCTGCGCTTATCCGGTACTATCGTCTTGAGT 1020
Db 961 TGACGAAACCCCGCTTACGCGCGGCTGCGCTTATCCGGTACTATCGTCTTGAGT 1020
Qy 1021 CCNACCCGTTAAGACACAGCTTATCGCCTGAGTGTAGTTCGCTCCAAGCTGGGCTGTG 1080
Db 1021 CCNACCCGTTAAGACACAGCTTATCGCCTGAGTGTAGTTCGCTCCAAGCTGGGCTGTG 1080
Qy 1081 GAGCAGGATGTAGTGGGCTGTACAGAGTTCCTTGAAGTGGTGGGCTAATACGGGTACA 1140
Db 1081 GAGCAGGATGTAGTGGGCTGTACAGAGTTCCTTGAAGTGGTGGGCTAATACGGGTACA 1140
Qy 1141 CTAGAAGGACAGTATTTGGTATCTGCGCTCTGCTGAAGCCAGTACCTTCGGAAGAAGAG 1200
Db 1141 CTAGAAGGACAGTATTTGGTATCTGCGCTCTGCTGAAGCCAGTACCTTCGGAAGAAGAG 1200
Qy 1201 TTGCTAGCTCTTGCATCGGCAAAACCAACCGCTGCTGAGCGGTGTTTTTTTGTGCA 1260
Db 1201 TTGCTAGCTCTTGCATCGGCAAAACCAACCGCTGCTGAGCGGTGTTTTTTTGTGCA 1260
Qy 1261 AGCAGCAGATTACGCGCAGAAAAAAGGATCTCGGGGATCCGGAGAGCTCCCAACCGCTT 1320

Db 1261 AGCAGCAGATTACCGCAGAAAAAGGATCTGGGGATCGGAGAGCTCCCAACGGTT 1320
QY 1321 GGATGATGATGAGGAAAGGAGGTAAGATCTGTAATGATAGCAGGAACTTTGAAGA 1380
Db 1321 GGATGATGATGAGGAAAGGAGGTAAGATCTGTAATGATAGCAGGAACTTTGAAGA 1380
QY 1381 CTCAGTGACTCAGTGAGTAATAAGACTCAGTGACTTCTGATCGTCTGCTGCTGCACT 1440
Db 1381 CTCAGTGACTCAGTGAGTAATAAGACTCAGTGACTTCTGATCGTCTGCTGCTGCACT 1440
QY 1441 CTTTGTGTCCTCAAGAAAGCGCTTCTGCTCTGAGGAGGAGCCCTTCCCTGGAAGGT 1500
Db 1441 CTTTGTGTCCTCAAGAAAGCGCTTCTGCTCTGAGGAGGAGCCCTTCCCTGGAAGGT 1500
QY 1501 AAACTAAGGATGTCAGCAGAGAAATTTTCCACCATTTGCTGCTCAAGAGGAAC 1560
Db 1501 AAACTAAGGATGTCAGCAGAGAAATTTTCCACCATTTGCTGCTCAAGAGGAAC 1560
QY 1561 TGATGACTCAGTCTAGATGAGAGCAGTGAGGAGAGACAGAGACTCGAATTTCCGGA 1620
Db 1561 TGATGACTCAGTCTAGATGAGAGCAGTGAGGAGAGACAGAGACTCGAATTTCCGGA 1620
QY 1621 GCTATTTCAGTTTCTTTCCGTTTGTGCAATTTCACTTATGATACCGGCAATGCTTG 1680
Db 1621 GCTATTTCAGTTTCTTTCCGTTTGTGCAATTTCACTTATGATACCGGCAATGCTTG 1680
QY 1681 GTTGCTATTGGAATCCCTTAGGGATGCCCTCAACTGGCCCTATAAGGGCCAG 1740
Db 1681 GTTGCTATTGGAATCCCTTAGGGATGCCCTCAACTGGCCCTATAAGGGCCAG 1740
QY 1741 CCTGAGCTGAGAGGATCTGTCAGAGGATCAAGACAGCAGTGAGCCTCGCAGAGCTC 1800
Db 1741 CCTGAGCTGAGAGGATCTGTCAGAGGATCAAGACAGCAGTGAGCCTCGCAGAGCTC 1800
QY 1801 TCCACA 1807
Db 1801 TCCACA 1807

RESULT 2
US-09-242-202a-28
; Sequence 28, Application US/09242202a
; GENERAL INFORMATION:
; APPLICANT: Nelson, Edward L.
; TITLE OF INVENTION: NOVEL VECTOR FOR
; POLYNUCLEOTIDE VACCINES
; NUMBER OF SEQUENCES: 29
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: MORGAN & FINNEGAN, L.L.P.
; STREET: 345 PARK AVENUE
; CITY: NEW YORK
; STATE: NEW YORK
; COUNTRY: USA
; ZIP: 10154
; COMPUTER READABLE FORM:
; MEDIUM TYPE: DISKETTE, 3.5 INCH, 1.44 MB
; COMPUTER: IBM PC COMPATIBLE
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: MS WORD 97
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/242, 202a
; FILING DATE: 20-Apr-2000
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US60/023931
; FILING DATE: 14-AUG-1996
; ATTORNEY/AGENT INFORMATION:
; NAME: KATHRYN M. BROWN
; REGISTRATION NUMBER: 34556
; REFERENCE/DOCKET NUMBER: 2026-4236US1
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 758-4800

; TELEFAX: (212) 751-6849
; TELEX: 421792
; INFORMATION FOR SEQ ID NO: 28:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1807 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: unknown
; MOLECULE TYPE: cdna
; HYPOTHEICAL: NO
; ANTI-SENSE: NO
; SEQUENCE DESCRIPTION: SEQ ID NO: 28:
US-09-242-202a-28

Query Match 100.0%; Score 1807; DB 16; Length 1807;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1807; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGTACCTGCCACCATTGCGCGGATTTCTTATCACTGATAGTTCGTCGACATATATGTT 60
Db 1 GGTACCTGCCACCATTGCGCGGATTTCTTATCACTGATAGTTCGTCGACATATATGTT 60

QY 61 TATCAGTGATAAGTGTCAGCATGACAAAGTTGACGCCGAATACAGTGATCCGTGCCG 120
Db 61 TATCAGTGATAAGTGTCAGCATGACAAAGTTGACGCCGAATACAGTGATCCGTGCCG 120

QY 121 CCCTGGACTGTTGAACGAGGTGCGGTAGACGGTCTCAGCACACGCAAACTGGCGGAC 180
Db 121 CCCTGGACTGTTGAACGAGGTGCGGTAGACGGTCTCAGCACACGCAAACTGGCGGAC 180

QY 181 GTTGGGGTGCAGCAGCGCGCTTTACTTGGCACTTCAGGAACAAGCGGCGCTTAAGG 240
Db 181 GTTGGGGTGCAGCAGCGCGCTTTACTTGGCACTTCAGGAACAAGCGGCGCTTAAGG 240

QY 241 GCCATATGTTGAGTGATGCTTGCACCCAGCGGGGATGGGGAGACCTCTAGTCAGAG 300
Db 241 GCCATATGTTGAGTGATGCTTGCACCCAGCGGGGATGGGGAGACCTCTAGTCAGAG 300

QY 301 CCCCAGGAGCAGCAGGCAATGCCGCTTCCCTGCGAGGATGAGTAGTGCGCTC 360
Db 301 CCCCAGGAGCAGCAGGCAATGCCGCTTCCCTGCGAGGATGAGTAGTGCGCTC 360

QY 361 TCCTGGCCCTGGAAGTTGCCACTCCAGTCCCGCCACAGCCTTGTCTAATAAATAAGTT 420
Db 361 TCCTGGCCCTGGAAGTTGCCACTCCAGTCCCGCCACAGCCTTGTCTAATAAATAAGTT 420

QY 421 GCATCATTTTGTCTGACTAGTGCTCTCTATAATATATATATATATATATATATAT 480
Db 421 GCATCATTTTGTCTGACTAGTGCTCTCTATAATATATATATATATATATATATAT 480

QY 481 CGGACTTTTGAAGTGATGTTGGGGGAGGATTCGAACCTTCGAGTCGATGACGGC 540
Db 481 CGGACTTTTGAAGTGATGTTGGGGGAGGATTCGAACCTTCGAGTCGATGACGGC 540

QY 541 AGATTAGAGTCTGCTCCCTTTGGCGCTCGGGAACCCACCGGTAATGCTTTTACT 600
Db 541 AGATTAGAGTCTGCTCCCTTTGGCGCTCGGGAACCCACCGGTAATGCTTTTACT 600

QY 601 GGCCTGCTCCCTTATCGGGAAGCGGCGCATCATATCAAAATGACGCCGCTCTAAAGT 660
Db 601 GGCCTGCTCCCTTATCGGGAAGCGGCGCATCATATCAAAATGACGCCGCTCTAAAGT 660

QY 661 GTTACGTTGAGAAAGAAATTCCTGAGCGCCCGGCTTGTGCTGGCGTTTTCATAGGCTCC 720
Db 661 GTTACGTTGAGAAAGAAATTCCTGAGCGCCCGGCTTGTGCTGGCGTTTTCATAGGCTCC 720

QY 721 GCCCCCTGAGGAGCATCAAAAATCGACGCTCAAGTCAGAGTGCGGGAACCGCAG 780
Db 721 GCCCCCTGAGGAGCATCAAAAATCGACGCTCAAGTCAGAGTGCGGGAACCGCAG 780

QY 781 GACTATAAAGATACAGGCGTTTCCCTCGGAAGCTCCCTGCTGCGCTCTCTCTGTCGGA 840
Db 781 GACTATAAAGATACAGGCGTTTCCCTCGGAAGCTCCCTGCTGCGCTCTCTCTGTCGGA 840

Db 781 GACTATAAGATACACAGCGTTTCCCTTGGAGCTCCCTCGTGCCTCTCTCTCCGA 840
QY 841 CCCTGCCGCTTACCGGATACCTGTCCGCTTCTCCCTTCGGGAAGCGTGGCTTCTC 900
Db 841 CCCTGCCGCTTACCGGATACCTGTCCGCTTCTCCCTTCGGGAAGCGTGGCTTCTC 900
QY 901 AATGCTCAGCTGTAGGTATCTCAGTTCGCTGTAGTTCGCTTCCCAAGCTGGCTGTG 960
Db 901 AATGCTCAGCTGTAGGTATCTCAGTTCGCTGTAGTTCGCTTCCCAAGCTGGCTGTG 960
QY 961 TGCACGAACCCCGCTTACGCGGCTGAGCGCTGCGCTTATCCGCTAACTATCGTCTGAGT 1020
Db 961 TGCACGAACCCCGCTTACGCGGCTGAGCGCTGCGCTTATCCGCTAACTATCGTCTGAGT 1020
QY 1021 CCAACCCGGTAAAGACAGCTATTCGCCACTGGCAGCAGCCACTGGTACAGATTAGCA 1080
Db 1021 CCAACCCGGTAAAGACAGCTATTCGCCACTGGCAGCAGCCACTGGTACAGATTAGCA 1080
QY 1081 GACGAGGTATGTAGCGGTGCTACAGAGTCTTGAAGTGGTGGCTAACTACGGGTACA 1140
Db 1081 GACGAGGTATGTAGCGGTGCTACAGAGTCTTGAAGTGGTGGCTAACTACGGGTACA 1140
QY 1141 CTAGAAGGACAGTATTGTGTATCTGCGCTCTGCTGAAGCCAGTTACCTTCGGAAGAG 1200
Db 1141 CTAGAAGGACAGTATTGTGTATCTGCGCTCTGCTGAAGCCAGTTACCTTCGGAAGAG 1200
QY 1201 TTGGTAGCTCTTATCGCGGCAACCAACACCCCTGTGTAGCGGTGCTTTTGTGTTGCA 1260
Db 1201 TTGGTAGCTCTTATCGCGGCAACCAACACCCCTGTGTAGCGGTGCTTTTGTGTTGCA 1260
QY 1261 AGCAGCAGATTACGCGCAGAAAAAGAGTCTGGGGATCCGAGAGCTCCCAACCGCTT 1320
Db 1261 AGCAGCAGATTACGCGCAGAAAAAGAGTCTGGGGATCCGAGAGCTCCCAACCGCTT 1320
QY 1321 GGATGATGATGAGGAAAGGAGTAAATCTGTAATGAATAGCAGGAACTTTGAGA 1380
Db 1321 GGATGATGATGAGGAAAGGAGTAAATCTGTAATGAATAGCAGGAACTTTGAGA 1380
QY 1381 CTCAGTCACTCAGTCACTAAAGACTCACTGACTTCTGATCTGAGGAGCCCTTCCCTGGAGGT 1440
Db 1381 CTCAGTCACTCAGTCACTAAAGACTCACTGACTTCTGATCTGAGGAGCCCTTCCCTGGAGGT 1440
QY 1441 CCTTGTGTTCCCAAGAGCGCTTCTGCTCTGAGGAGCCCTTCCCTGGAGGT 1500
Db 1441 CCTTGTGTTCCCAAGAGCGCTTCTGCTCTGAGGAGCCCTTCCCTGGAGGT 1500
QY 1501 AAACTAAGATGTCAGCAGAGAAATTTTCCACCATTTGCTGTGCTTCAAGAGGAAC 1560
Db 1501 AAACTAAGATGTCAGCAGAGAAATTTTCCACCATTTGCTGTGCTTCAAGAGGAAC 1560
QY 1561 TGATCAGTCACTCTAGATGAGCAGAGTCTGAGGAGAGCAGAGCTCGAATTTCCGGA 1620
Db 1561 TGATCAGTCACTCTAGATGAGCAGAGTCTGAGGAGAGCAGAGCTCGAATTTCCGGA 1620
QY 1621 GCTATTTTCAGTCTTCTTTCGTTTTTGTGCAATTTCACTTATGATACCGGCAATGCTG 1680
Db 1621 GCTATTTTCAGTCTTCTTTCGTTTTTGTGCAATTTCACTTATGATACCGGCAATGCTG 1680
QY 1681 GTTGCTATTTTGGAACTCCCTTAGGGGATGCCCCCTCACTGCTTATTAAGGGCCAG 1740
Db 1681 GTTGCTATTTTGGAACTCCCTTAGGGGATGCCCCCTCACTGCTTATTAAGGGCCAG 1740
QY 1741 CCTGAGCTGCAGAGATTCTTCGAGAGGATCAAGACAGCAGCTGGACCTCGCAGAGCTC 1800
Db 1741 CCTGAGCTGCAGAGATTCTTCGAGAGGATCAAGACAGCAGCTGGACCTCGCAGAGCTC 1800
QY 1801 TCCACACA 1807
Db 1801 TCCACACA 1807

Sequence 29, Application US/09242202
GENERAL INFORMATION:
APPLICANT: Nelson, Edward L.
Nelson, Peter J.
TITLE OF INVENTION: NOVEL VECTOR FOR POLYNUCLEOTIDE VACCINES
NUMBER OF SEQUENCES: 29
CORRESPONDENCE ADDRESS:
ADDRESSEE: MORGAN & FINNEGAN, L.L.P.
STREET: 345 PARK AVENUE
CITY: NEW YORK
STATE: NEW YORK
COUNTRY: USA
ZIP: 10154
COMPUTER READABLE FORM:
MEDIUM TYPE: DISKETTE, 3.5 INCH, 1.44 MB
COMPUTER: IBM PC COMPATIBLE
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: MS WORD 97
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/242.202
FILING DATE: 01-Nov-1999
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US60/023931
FILING DATE: 14-AUG-1996
ATTORNEY/AGENT INFORMATION:
NAME: KATHRYN M. BROWN
REGISTRATION NUMBER: 34556
REFERENCE/DOCKET NUMBER: 2026-4236US1
TELECOMMUNICATION INFORMATION:
TELEPHONE: (212) 758-4800
TELEFAX: (212) 751-6849
TELEX: 421792
INFORMATION FOR SEQ ID NO: 29:
SEQUENCE CHARACTERISTICS:
LENGTH: 2308 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: unknown
MOLECULE TYPE: CDNA
HYPOTHETICAL: NO
ANTI-SENSE: NO
SEQUENCE DESCRIPTION: SEQ ID NO: 29:
US-09-242-202-29

Query Match 87.6%; Score 1583.2; DB 16; Length 2308;
Best Local Similarity 99.8%; Pred. No. 0;
Matches 1585; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
QY 220 GACAAAGCGGGCCCTTAAGGGCCATATGGTGAAGTGGATGCCTTGACCCAGCGGGGAT 279
Db 721 GTACAAAGTGAAGCGCTTAAGGGCCATATGGTGAAGTGGATGCCTTGACCCAGCGGGGAT 780
QY 280 GGGGAGACCTGTAGTCAGAGCCCGGGCAGCAGCAAGCAATGCCCTCTCCCTCCCTGC 339
Db 781 GGGGAGACCTGTAGTCAGAGCCCGGGCAGCAGCAAGCAATGCCCTCTCCCTCCCTGC 840
QY 340 AGGATGAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGT 399
Db 841 AGGATGAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGT 900
QY 400 TTGCTCTAATAAATAAGTGGATCATTTTGTCTGACTAGGTGCTCTCTAATAATAATAT 459
Db 901 TTGCTCTAATAAATAAGTGGATCATTTTGTCTGACTAGGTGCTCTCTAATAATAATAT 960
QY 460 AAGCTTGATGAAATTTCTTTTCGAGCTTTTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGT 519
Db 961 AAGCTTGATGAAATTTCTTTTCGAGCTTTTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGT 1020
QY 520 ACCTTCGAAGTCGATGACGCGCAGATTTAGAGTCTGTCTCTTGGCGCGCTCGGGAACCCC 579
Db 1021 ACCTTCGAAGTCGATGACGCGCAGATTTAGAGTCTGTCTCTTGGCGCGCTCGGGAACCCC 1080

QY	580	ACCACGGGTAATGCTTTTAACTGGCCTGCTCCCTTATCGGGAAGCGGGCGCATCATATCA	639
DB	1081	ACCACGGGTAATGCTTTTAACTGGCCTGCTCCCTTATCGGGAAGCGGGCGCATCATATCA	1140
QY	640	AATGACGCGCCGCTGTAAGTGTACGTTTGAGAAAGAAATTCCTGCAGCCGCCGCGTTGC	699
DB	1141	AATGACGCGCCGCTGTAAGTGTACGTTTGAGAAAGAAATTCCTGCAGCCGCCGCGTTGC	1200
QY	700	TGCGGTTTTTCCATAGGCTCCGCCCTTGACGAGCATCAAAAATCGAGCGTCAAGTC	759
DB	1201	TGCGGTTTTTCCATAGGCTCCGCCCTTGACGAGCATCAAAAATCGAGCGTCAAGTC	1260
QY	760	AGAGGTGGGAAACCCGACGAGCACTATAAGATACCAAGCTTTCCCCCTGGAAAGTCCC	819
DB	1261	AGAGGTGGGAAACCCGACGAGCACTATAAGATACCAAGCTTTCCCCCTGGAAAGTCCC	1320
QY	820	TGCTGCGCTCTCCTGTTTCGACCCCTGCCGCTTACCGGATACCTCTGCGCCTTTCTCCCTT	879
DB	1321	TGCTGCGCTCTCCTGTTTCGACCCCTGCCGCTTACCGGATACCTCTGCGCCTTTCTCCCTT	1380
QY	880	CGGGAACGCTGGCGCTTTTCAATGCTCACGCTGTAGTATCTCAGTTTCGGTGTAGTGTG	939
DB	1381	CGGGAACGCTGGCGCTTTTCAATGCTCACGCTGTAGTATCTCAGTTTCGGTGTAGTGTG	1440
QY	940	TTGCTTCCAAAGCTGGCGTGTGCAGCAACCCCGGTTACGCCGACCGCTGCGCCTTAT	999
DB	1441	TTGCTTCCAAAGCTGGCGTGTGCAGCAACCCCGGTTACGCCGACCGCTGCGCCTTAT	1500
QY	1000	CCGGTAACATCTCTTTGAGTCCAAACCCGGTAAGACACGACTTATCGCACTGGCAGCAG	1059
DB	1501	CCGGTAACATCTCTTTGAGTCCAAACCCGGTAAGACACGACTTATCGCACTGGCAGCAG	1560
QY	1060	CCACTGGTAACAGGATTACGAGAGCGAGTATGTAGGCGGTGCTACAGAGTTCTTGAAGT	1119
DB	1561	CCACTGGTAACAGGATTACGAGAGCGAGTATGTAGGCGGTGCTACAGAGTTCTTGAAGT	1620
QY	1120	GGTGGCCTAACATACGGCTACACTAGAAGACAGTATTGGTATCTGGCGCTCTGCTGAAGC	1179
DB	1621	GGTGGCCTAACATACGGCTACACTAGAAGACAGTATTGGTATCTGGCGCTCTGCTGAAGC	1680
QY	1180	CAGTTACCTTCGGAAAAAGATTGGTAGCTCTTGATCCGGCAAAACAAACACCGCTGGTA	1239
DB	1681	CAGTTACCTTCGGAAAAAGATTGGTAGCTCTTGATCCGGCAAAACAAACACCGCTGGTA	1740
QY	1240	GCGGTGGTTTTTTTGTTCGAAGCAGCAGATPACGCGCAGAAAAAAGGATCTCGGGGAT	1299
DB	1741	GCGGTGGTTTTTTTGTTCGAAGCAGCAGATPACGCGCAGAAAAAAGGATCTCGGGGAT	1800
QY	1300	CCGGAGAGCTCCCAAGCGCTTGGATGCATGGATGAGGAAAGGAGGTAAAGATCTGTAAATG	1359
DB	1801	CCGGAGAGCTCCCAAGCGCTTGGATGCATGGATGAGGAAAGGAGGTAAAGATCTGTAAATG	1860
QY	1360	AATAAGCAGGAACCTTTGAAAGCTCAGTGACTCAGTGAGTAAATAAGACTCAGTGACTTCT	1419
DB	1861	AATAAGCAGGAACCTTTGAAAGCTCAGTGACTCAGTGAGTAAATAAGACTCAGTGACTTCT	1920
QY	1420	GATCTCTGCTTAAGTCCGCTCTCTTGTGTCGAAGAAAGCGGCTTCCTGCTCTCTGAGG	1479
DB	1921	GATCTCTGCTTAAGTCCGCTCTCTTGTGTCGAAGAAAGCGGCTTCCTGCTCTCTGAGG	1980
QY	1480	AGGACCCCTTCCTCGAAGGTAAACTAAGGATGTACGACAGAAATTTTCCACCAATG	1539
DB	1981	AGGACCCCTTCCTCGAAGGTAAACTAAGGATGTACGACAGAAATTTTCCACCAATG	2040
QY	1540	GTGCTTTGGTCAAAAGAGAAACTGATGAGCTCACTCTAGATGAGAGCAGTGAGGGAGAG	1599
DB	2041	GTGCTTTGGTCAAAAGAGAAACTGATGAGCTCACTCTAGATGAGAGCAGTGAGGGAGAG	2100
QY	1500	ACAGAGACTCGAAATTTCCGGAGGCTATTTACGTTTTCCTTTCCGTTTGTGCAATTTCACT	1559
DB	2101	ACAGAGACTCGAAATTTCCGGAGGCTATTTACGTTTTCCTTTCCGTTTGTGCAATTTCACT	2160

QY	1660	TATGATACCGGCAATGCTTGGTTGCTATTTTGGAAACTCCCTTAGGGGATGCCCTCA	1719
Db	2161	TATGATACCGGCAATGCTTGGTTGCTATTTTGGAAACTCCCTTAGGGGATGCCCTCA	2220
QY	1720	ACTGGCCCTATAAAGGCCAGCGCTGAGCTGCAGAGGATTCCTGCAGAGGATCAAGACAGC	1779
Db	2221	ACTGGCCCTATAAAGGCCAGCGCTGAGCTGCAGAGGATTCCTGCAGAGGATCAAGACAGC	2280
QY	1780	ACGTGGACCTCGCACAGCCTCTCCACACA	1807
Db	2281	ACGTGGACCTCGCACAGCCTCTCCACACA	2308
RESULT 4			
US-09-242-202A-29			
; Sequence 29, Application US/09242202A			
; GENERAL INFORMATION:			
; APPLICANT: Nelson, Edward L.			
; Nelson, Peter J.			
; TITLE OF INVENTION: NOVEL VECTOR FOR POLYNUCLEOTIDE VACCINES			
; NUMBER OF SEQUENCES: 29			
; CORRESPONDENCE ADDRESS:			
; ADDRESSEE: MORGAN & FINNEGAN, L.L.P.			
; STREET: 345 PARK AVENUE			
; CITY: NEW YORK			
; STATE: NEW YORK			
; COUNTRY: USA			
; ZIP: 10154			
; COMPUTER READABLE FORM:			
; MEDIUM TYPE: DISKETTE, 3.5 INCH, 1.44 MB			
; COMPUTER: IBM PC COMPATIBLE			
; OPERATING SYSTEM: PC-DOS/MS-DOS			
; SOFTWARE: MS WORD 97			
; CURRENT APPLICATION DATA:			
; APPLICATION NUMBER: US/09/242,202A			
; FILING DATE: 20-Apr-2000			
; PRIOR APPLICATION DATA:			
; APPLICATION NUMBER: US60/023931			
; FILING DATE: 14-AUG-1996			
; ATTORNEY/AGENT INFORMATION:			
; NAME: KATHRYN M. BROWN			
; REGISTRATION NUMBER: 34556			
; REFERENCE/DOCKET NUMBER: 2026-4236US1			
; TELECOMMUNICATION INFORMATION:			
; TELEPHONE: (212) 758-4800			
; TELEFAX: (212) 751-6849			
; TELEX: 421792			
; INFORMATION FOR SEQ ID NO: 29:			
; SEQUENCE CHARACTERISTICS:			
; LENGTH: 2308 base pairs			
; TYPE: nucleic acid			
; STRANDEDNESS: single			
; TOPOLOGY: unknown			
; MOLECULE TYPE: CDNA			
; HYPOTHETICAL: NO			
; ANTI-SENSE: NO			
; SEQUENCE DESCRIPTION: SEQ ID NO: 29:			
US-09-242-202A-29			

Query Match	87.6%	Score 1583.2;	DB 16;	Length 2308;
Best Local Similarity	99.8%	Pred. No. 0;		
Matches 1585; Conservative	0;	Mismatches	3;	Indels 0; Gaps 0;

Oy	220	GAACAAGCGGCGCCTTAAAGGCCATATGTGTAGTGTGATGCCCTTGACCCACAGCGGGGAT	279
Db	721	GTACAAGTGAAGCGCCTTTAAAGGCCATATGTGTAGTGTGATGCCCTTGACCCACAGCGGGGAT	780
Oy	280	GGGGGAGACCTGTAGTCAGAGCCCCGGGGGAGACACAGGCCCAATGCCGTCTTCCCTCGC	339
Db	781	GGGGGAGACCTGTAGTCAGAGCCCCGGGGGAGACACAGGCCCAATGCCGTCTTCCCTCGC	840

QY 340 AGGATGAGTAGTGGCTCTCTCTGCGCCCTGGAGTTGCCACTCCAGTCCGCCACAGCC 399
DB 841 AGGATGAGTAGTGGCTCTCTCTGCGCCCTGGAGTTGCCACTCCAGTCCGCCACAGCC 900
QY 400 TTGTCCTAATAAATAAAGTTAGTTCATATTTTCTGCTAGTGTCTCTCTATAATATAT 459
DB 901 TTGTCCTAATAAATAAAGTTAGTTCATATTTTCTGCTAGTGTCTCTATAATATAT 960
QY 460 AAGCTTGATATCGAATCTTTTCGGACATTTTGAAGTGTAGTGTGGGGGGAAGATTCGA 519
DB 961 AAGCTTGATATCGAATCTTTTCGGACATTTTGAAGTGTAGTGTGGGGGGAAGATTCGA 1020
QY 520 ACCTTGGAAGTGTAGCGAGATTTAGAGTCTGCTCCCTTTGGCGCTCGGGAACCCC 579
DB 1021 ACCTTGGAAGTGTAGCGAGATTTAGAGTCTGCTCCCTTTGGCGCTCGGGAACCCC 1080
QY 580 ACCAGGGTAAATGCTTTTACTGGCTGCTCCCTTTATCGGGAAGCGGGCGCATCATATCA 639
DB 1081 ACCAGGGTAAATGCTTTTACTGGCTGCTCCCTTTATCGGGAAGCGGGCGCATCATATCA 1140
QY 640 AATAGCGCGCCCTGTAAAGTGTAGTTGAGAAAGATTTCTGACGCGCCGCGTTCG 699
DB 1141 AATAGCGCGCCCTGTAAAGTGTAGTTGAGAAAGATTTCTGACGCGCCGCGTTCG 1200
QY 700 TGGCGTTTTCATAGGCTCCGCCCTGACGAGCATCAAAAAATGAGCTCAAGTC 759
DB 1201 TGGCGTTTTCATAGGCTCCGCCCTGACGAGCATCAAAAAATGAGCTCAAGTC 1260
QY 760 AGAGTGGCGAAACCGCAGGACTATAAAGATACAGCGCTTCCCGTGAAGCTGCC 819
DB 1261 AGAGTGGCGAAACCGCAGGACTATAAAGATACAGCGCTTCCCGTGAAGCTGCC 1320
QY 820 TGTGCGCTCTCTGTTCCGACCCCTGCGCTTACCGGATACCTGCGCTTCTCCCTT 879
DB 1321 TGTGCGCTCTCTGTTCCGACCCCTGCGCTTACCGGATACCTGCGCTTCTCCCTT 1380
QY 880 CGGGAAGCTGGCGCTTCTCAATGCTCAGCTGTAGTATCTCAGTTCCGCTAGGTCG 939
DB 1381 CGGGAAGCTGGCGCTTCTCAATGCTCAGCTGTAGTATCTCAGTTCCGCTAGGTCG 1440
QY 940 TTCGCTCCAAGCTGGCGCTGTGTCAGCAACCCCGCTTACGCCCGCAGCTGCGCTTAT 999
DB 1441 TTCGCTCCAAGCTGGCGCTGTGTCAGCAACCCCGCTTACGCCCGCAGCTGCGCTTAT 1500
QY 1000 CCGGTAACTATCTCTGAGTCCAAACCGGTAAAGACACACTTATCCGCTGCGAGCAG 1059
DB 1501 CCGGTAACTATCTCTGAGTCCAAACCGGTAAAGACACACTTATCCGCTGCGAGCAG 1560
QY 1060 CCACTGTAACAGGATTAGCAGACGAGGTATGTCGCGGTGTACAGAGTCTTGAAGT 1119
DB 1561 CCACTGTAACAGGATTAGCAGACGAGGTATGTCGCGGTGTACAGAGTCTTGAAGT 1620
QY 1120 GGTGGCCCTAAGCTGCTACACTAGAGGACAGTATTTGGTATCTCGCTCTCTGAAGC 1179
DB 1621 GGTGGCCCTAAGCTGCTACACTAGAGGACAGTATTTGGTATCTCGCTCTCTGAAGC 1680
QY 1180 CAGTTACTTTGGAAAAAGAGTTGGTACCTTGTATCCGCAACCAACCAACCGCTGGTA 1239
DB 1581 CAGTTACTTTGGAAAAAGAGTTGGTACCTTGTATCCGCAACCAACCAACCGCTGGTA 1740
QY 1240 GCGGTGTTTTTTTGGTGGAGCAGATTTAGCGCCAGAAAAAAGATCTGGGGGAT 1299
DB 1741 GCGGTGTTTTTTTGGTGGAGCAGATTTAGCGCCAGAAAAAAGATCTGGGGGAT 1800
QY 1300 CCGGAGAGCTCCCAACCGTTGGATGATGATGAGGAAAGAGGTAAAGTCTGTAATG 1359
DB 1801 CCGGAGAGCTCCCAACCGTTGGATGATGATGAGGAAAGAGGTAAAGTCTGTAATG 1860
QY 1360 AATAAGCAGGAACCTTGAAGACTCAGTCACTAGTGTAGTAAATAAGACTCAGTCACTTCT 1419
DB 1861 AATAAGCAGGAACCTTGAAGACTCAGTCACTAGTGTAGTAAATAAGACTCAGTCACTTCT 1920
QY 1420 GATCTGTCTCACTGCCACTCTTGTGTGTCACCAAGAAAGCGGCTTCTCTCTCTGAGG 1479

DB 1921 GATCTGTCTCACTGCCACTCTTGTGTGTCACCAAGAAAGCGGCTTCTCTCTCTGAGG 1980
QY 1480 AGGACCCCTTCCCTGGAAGCTAAAGTAAAGTGTAGGATGTGAGAGAGAAATTTTCCACCATG 1539
DB 1981 AGGACCCCTTCCCTGGAAGCTAAAGTAAAGTGTAGGATGTGAGAGAGAAATTTTCCACCATG 2040
QY 1540 GTGCTTGGTCAAAAGAGAAACTGATGAGCTCACTCTAGATGAGAGAGCAGTGGAGGAG 1599
DB 2041 GTGCTTGGTCAAAAGAGAAACTGATGAGCTCACTCTAGATGAGAGAGCAGTGGAGGAG 2100
QY 1600 ACAGAGACTCGAATTTCCGAGCTATTTCACTTTTCCGTTTGTGCAATTTCACT 1659
DB 2101 ACAGAGACTCGAATTTCCGAGCTATTTCACTTTTCCGTTTGTGCAATTTCACT 2160
QY 1660 TATGATACCGGCCAATCTTGGTGTCTATTTTGGAACTCCCTTAGGGGATGCCCTCA 1719
DB 2161 TATGATACCGGCCAATCTTGGTGTCTATTTTGGAACTCCCTTAGGGGATGCCCTCA 2220
QY 1720 ACTGGCCCTATAAAGGCGCAGCTGAGCTGAGAGGATTTCTGAGAGGATCAAGACAGC 1779
DB 2221 ACTGGCCCTATAAAGGCGCAGCTGAGCTGAGAGGATTTCTGAGAGGATCAAGACAGC 2280
QY 1780 ACSTGGACCTCGCACAGCTCTCCACCA 1807
DB 2281 ACSTGGACCTCGCACAGCTCTCCACCA 2308

RESULT 5
US-09-242-202-27
; Sequence 27, Application US/09242202
; GENERAL INFORMATION:
; APPLICANT: Nelson, Edward L.
; Nelson, Peter J.
; TITLE OF INVENTION: NOVEL VECTOR FOR
; POLYNUCLEOTIDE VACCINES
; NUMBER OF SEQUENCES: 29
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: MORGAN & FINNEGAN, L.L.P.
; STREET: 345 PARK AVENUE
; CITY: NEW YORK
; STATE: NEW YORK
; COUNTRY: USA
; ZIP: 10154
; COMPUTER READABLE FORM:
; MEDIUM TYPE: DISKETTE, 3.5 INCH, 1.44 MB
; COMPUTER: IBM PC COMPATIBLE
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: MS WORD 97
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/242,202
; FILING DATE: 01-NOV-1999
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US60/023931
; FILING DATE: 14-AUG-1996
; ATTORNEY/AGENT INFORMATION:
; NAME: KATHRYN M. BROWN
; REGISTRATION NUMBER: 34556
; REFERENCE/DOCKET NUMBER: 2026-4236US1
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 758-4800
; TELEFAX: (212) 751-6849
; TELEX: 421792
; INFORMATION FOR SEQ ID NO: 27:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1547 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: unknown
; MOLECULE TYPE: cDNA
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
; SEQUENCE DESCRIPTION: SEQ ID NO: 27:

US-09-242-202-27

Query Match 72.5%; Score 1310.4; DB 16; Length 1547;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 1311; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 GGTACCTGCCACCATGGCGGGATCTTTATCACTGATAAGTTGGTGACATATATGTT 60
DB 1 GGTACCTGCCACCATGGCGGGATCTTTATCACTGATAAGTTGGTGACATATATGTT 60
QY 61 TATCAGTGATAAAGTGTCAAGCATGACAAAGTTGCGAGCCGAATACAGTGCCTGCGG 120
DB 61 TATCAGTGATAAAGTGTCAAGCATGACAAAGTTGCGAGCCGAATACAGTGCCTGCGG 120
QY 121 CCCTGGACTGTTGAACGAGGTGCGGTAGACGCTGACGACGCAAACTGGCGGAACG 180
DB 121 CCCTGGACTGTTGAACGAGGTGCGGTAGACGCTGACGACGCAAACTGGCGGAACG 180
QY 181 GTTGGGGTGCACGACGCGCGCTTTACTGGCACTTCAGGAACAGCGGCGCTTAAGG 240
DB 181 GTTGGGGTGCACGACGCGCGCTTTACTGGCACTTCAGGAACAGCGGCGCTTAAGG 240
QY 241 GCCATATGTTGAGTGGATGCTTTGACCCAGCGGGGATGGGGGAGACCTGTAGTCAGAG 300
DB 241 GCCATATGTTGAGTGGATGCTTTGACCCAGCGGGGATGGGGGAGACCTGTAGTCAGAG 300
QY 301 CCCCCGGGACGACAGCCCAATCCCGTCTTCCCTGCAGGATGAGTAGTGCCCTC 360
DB 301 CCCCCGGGACGACAGCCCAATCCCGTCTTCCCTGCAGGATGAGTAGTGCCCTC 360
QY 361 TCCTGGCCCTGGAAGTGGCCACTCCAGTGCACACGCGCTTGTCTAATAAATAAAGTT 420
DB 361 TCCTGGCCCTGGAAGTGGCCACTCCAGTGCACACGCGCTTGTCTAATAAATAAAGTT 420
QY 421 GCATCATTTTGTCTGACTAGTGTCTCTATAATATATAAGCTTGATATCGAATCTTT 480
DB 421 GCATCATTTTGTCTGACTAGTGTCTCTATAATATATAAGCTTGATATCGAATCTTT 480
QY 481 CGGACITTTTGAAGATGATGTGTGGGGAGGATTCGAACCTTCGAGTCGAGCGC 540
DB 481 CGGACITTTTGAAGATGATGTGTGGGGAGGATTCGAACCTTCGAGTCGAGCGC 540
QY 541 AGATTTAGACTGTCTCCCTTTGGCGCTCGGAAACCCACCGGTAATGCTTTTACT 600
DB 541 AGATTTAGACTGTCTCCCTTTGGCGCTCGGAAACCCACCGGTAATGCTTTTACT 600
QY 601 GGCTGCTCCCTTTATCGGGAAGCGGGCGCATCATATCAATGACGCGCGCTGTAAAGT 660
DB 601 GGCTGCTCCCTTTATCGGGAAGCGGGCGCATCATATCAATGACGCGCGCTGTAAAGT 660
QY 661 GTTACGTTGAGAAGATTCCTGACGCGCGCGCTTGTGGCTTTTCCATAGGCTCC 720
DB 661 GTTACGTTGAGAAGATTCCTGACGCGCGCGCTTGTGGCTTTTCCATAGGCTCC 720
QY 721 GCCCCCTGACGAGCATCAAAAATCGACGCTCAAGTCAGAGGTGCGGAAACCCGACAG 780
DB 721 GCCCCCTGACGAGCATCAAAAATCGACGCTCAAGTCAGAGGTGCGGAAACCCGACAG 780
QY 781 GACTATAAAGATACAGGCGCTTTCCCTCGGAAGCTCCCTGCGCTCCCTGTTCCGA 840
DB 781 GACTATAAAGATACAGGCGCTTTCCCTCGGAAGCTCCCTGCGCTCCCTGTTCCGA 840
QY 841 CCCTGGCGCTTACGGATACCTGCGGCTTCTCCCTTCGGGAAGGCTGGCGCTTCTC 900
DB 841 CCCTGGCGCTTACGGATACCTGCGGCTTCTCCCTTCGGGAAGGCTGGCGCTTCTC 900
QY 901 AATGCTCACGCTGTAGTATCTCAGTTCGCTGAGTGTAGGTCGCTCAAGCTGGCGTGTG 960
DB 901 AATGCTCACGCTGTAGTATCTCAGTTCGCTGAGTGTAGGTCGCTCAAGCTGGCGTGTG 960
QY 961 TGCACGAACCCCGTTTACGCGGACGCTGCGCTTATCCGCTTAACCTATCGTCTTGA 1020
DB 961 TGCACGAACCCCGTTTACGCGGACGCTGCGCTTATCCGCTTAACCTATCGTCTTGA 1020

DB 961 TGCACGAACCCCGTTTACGCGGACGCTGCGCTTATCCGCTTAACCTATCGTCTTGA 1020
QY 1021 CCAACCCGGTAAACACAGCACTTATCGCCACTGGCAGCAGCCACTGGTAAACAGATTAGCA 1080
DB 1021 CCAACCCGGTAAACACAGCACTTATCGCCACTGGCAGCAGCCACTGGTAAACAGATTAGCA 1080
QY 1081 GAGCGAGTATGTAGGCGGTGCTACAGAGTTCTTGAAGTGGTGCCTTAACCTACGGCTTACA 1140
DB 1081 GAGCGAGTATGTAGGCGGTGCTACAGAGTTCTTGAAGTGGTGCCTTAACCTACGGCTTACA 1140
QY 1141 CTAGAGGACAGATTTGGTGTATCTGCGCTCTGCTGAAGCCAGTTACCTTCGGAAGAAAG 1200
DB 1141 CTAGAGGACAGATTTGGTGTATCTGCGCTCTGCTGAAGCCAGTTACCTTCGGAAGAAAG 1200
QY 1201 TTGGTAGCTCTTGATCCGGCAACAAACACCAACCGCTGCTGAGCGTGGTGGTTTGGTTTGA 1260
DB 1201 TTGGTAGCTCTTGATCCGGCAACAAACACCAACCGCTGCTGAGCGTGGTGGTTTGGTTTGA 1260
QY 1261 AGCAGCAGATTACGCGCAGAAAAAGGATCTGGGGATCCGGAGAGCTCC 1312
DB 1261 AGCAGCAGATTACGCGCAGAAAAAGGATCTGGGGATCCGGAGAGCTCC 1312

RESULT 6

US-09-242-202a-27

; Sequence 27, Application US/09242202A

; GENERAL INFORMATION:

; APPLICANT: Nelson, Edward L.

; TITLE OF INVENTION: NOVEL VECTOR FOR

; POLYNUCLEOTIDE VACCINES

; NUMBER OF SEQUENCES: 29

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: MORGAN & FINNEGAN, L.L.P.

; STREET: 345 PARK AVENUE

; CITY: NEW YORK

; STATE: NEW YORK

; COUNTRY: USA

; ZIP: 10154

; COMPUTER READABLE FORM:

; MEDIUM TYPE: DISKETTE, 3.5 INCH, 1.44 MB

; COMPUTER: IBM PC COMPATIBLE

; OPERATING SYSTEM: PC-DOS/MS-DOS

; SOFTWARE: MS WORD 97

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/09/242,202A

; FILING DATE: 20-Apr-2000

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: US60/023931

; FILING DATE: 14-AUG-1996

; ATTORNEY/AGENT INFORMATION:

; NAME: KATHRYN M. BROWN

; REGISTRATION NUMBER: 34556

; REFERENCE/DOCKET NUMBER: 2026-4236US1

; TELECOMMUNICATION INFORMATION:

; TELEPHONE: (212) 758-4800

; TELEFAX: (212) 751-6849

; TELEX: 421792

; INFORMATION FOR SEQ ID NO: 27:

; SEQUENCE CHARACTERISTICS:

; LENGTH: 1547 base pairs

; TYPE: nucleic acid

; STRANDEDNESS: single

; TOPOLOGY: unknown

; MOLECULE TYPE: cDNA

; HYPOTHETICAL: No

; ANTI-SENSE: No

; SEQUENCE DESCRIPTION: SEQ ID NO: 27:

US-09-242-202a-27

Query Match 72.5%; Score 1310.4; DB 16; Length 1547;
Best Local Similarity 99.9%; Pred. No. 0;

QY 1020 TCCAAACCCGTAAGACACGACTTATGCCACTGCGCAGCAGCCACTGTAACAGGATTAGC 1079
 Db 5422 tcaaccccgtaagacacgacttatgccactgagcagcagccactgtaacagagattagc 5481
 QY 1080 AGACGAGGATATAGCGGTGTACAGAGTCTTGAAGTGGCTAACTACGGCTAC 1139
 Db 5482 agacgaggtatgtagcggtgtacagagttcttgaagtggtagcagcagcagcagcagc 5541
 QY 1140 ACTAGAGGACACTATTGTGTATCTGGCTCTGCTGAAGCCAGTTACCTTCGGAAGAAGA 1199
 Db 5542 accagaagagcagattatgtgtatctgctgctgctgctgaagccagttaccttcggaagaaga 5601
 QY 1200 GTTGGTAGCTCTTATCGCGCAACAAACACCGCTGGTGAAGTGGCTAACTACGGCTAC 1259
 Db 5602 gtggtagctcttgcgcaacaaacacccgctggtgaagccagttaccttcggaagaaga 5661
 QY 1260 AAGCAGCAGATTACGGCGCAGAAAAAAGGATCT 1292
 Db 5662 aagcagcagattacgcgcagaaaaaaggatct 5694

RESULT 8
 US-08-480-120-15
 ; Sequence 15, Application US/08480120
 ; GENERAL INFORMATION:
 ; APPLICANT: Joliffe, Linda K.
 ; APPLICANT: Zivin, Robert A.
 ; APPLICANT: Pulito, Virginia L.
 ; TITLE OF INVENTION: CDR-GRAFTED ANTI-TISSUE FACTOR
 ; TITLE OF INVENTION: ANTIBODIES AND METHODS OF USE THEREOF
 ; NUMBER OF SEQUENCES: 20
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: Scully, Scott, Murphy & Presser
 ; STREET: 400 Garden City Plaza
 ; CITY: Garden City
 ; STATE: New York
 ; COUNTRY: United States
 ; ZIP: 11530
 ; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: Floppy disk
 ; COMPUTER: IBM PC compatible
 ; OPERATING SYSTEM: PC-DOS/MS-DOS
 ; SOFTWARE: Patent In Release #1.0, Version #1.25
 ; CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: US/08/480,120
 ; FILING DATE: 07-JUN-1995
 ; CLASSIFICATION: 530
 ; ATTORNEY/AGENT INFORMATION:
 ; NAME: Digiglio, Frank S.
 ; REGISTRATION NUMBER: 31,346
 ; REFERENCE/DOCKET NUMBER: 9598
 ; TELECOMMUNICATION INFORMATION:
 ; TELEPHONE: (516) 742-4343
 ; TELEFAX: (516) 742-4366
 ; TELEX: 230 901 SANS UR
 ; INFORMATION FOR SEQ ID NO: 15:
 ; SEQUENCE CHARACTERISTICS:
 ; LENGTH: 7073 base pairs
 ; TYPE: nucleic acid
 ; STRANDEDNESS: double
 ; TOPOLOGY: linear
 ; MOLECULE TYPE: peptide
 ; FEATURE:
 ; NAME/KEY: CDS
 ; LOCATION: 61..717
 ; FEATURE:
 ; NAME/KEY: CDS
 ; LOCATION: 1111..1146
 ; FEATURE:
 ; NAME/KEY: CDS
 ; LOCATION: 1268..1594
 ; FEATURE:

NAME/KEY: CDS
 LOCATION: 1692..2012
 US-08-480-120-15
 Query Match 33.4%; Score 604.2; DB 8; Length 7073;
 Best Local Similarity 97.9%; Pred. No. 1.8e-167;
 Matches 612; Conservative 0; Mismatches 13; Indels 0; Gaps 0;
 QY 568 TGAAGAAGAAATTCCTGCAGCGCGCGCGCTTGCCTGCGCGCTTTTCCATAGGCTCCGCCCC 727
 Db 2856 TAAGGAGAGAGCGCTGCAGCTCGCGCGCGCTTGCCTGCGCGCTTTTCCATAGGCTCCGCCCC 2915
 QY 728 TGACGAGCATCAAAAAATCAGCTCAAGTCAGAGGTGGCGAAACCCGACAGGACTATA 787
 Db 2916 TGACGAGCATCAAAAAATCAGCTCAAGTCAGAGGTGGCGAAACCCGACAGGACTATA 2975
 QY 788 AAGATACCGAGCGTTTCCCGCTGGAAGCTCCCTGCTGCGCTCTCTCTCCGACCCGTC 847
 Db 2976 AAGATACCGAGCGTTTCCCGCTGGAAGCTCCCTGCTGCGCTCTCTCTCCGACCCGTC 3035
 QY 848 GCTTACCGGATACCTGTCGCGCTTTCCTCCCTTCGGGAAGCGTGGCGCTTCTCAATGCTC 907
 Db 3036 GCTTACCGGATACCTGTCGCGCTTTCCTCCCTTCGGGAAGCGTGGCGCTTCTCAATGCTC 3095
 QY 908 ACGCTGTAGGTATCTCAGTTCGGGTGTAGTTCGCTTCCAAAGCTGGGCTGTGTGCACGA 967
 Db 3096 ACGCTGTAGGTATCTCAGTTCGGGTGTAGTTCGCTTCCAAAGCTGGGCTGTGTGCACGA 3155
 QY 968 ACCCCCGCTTTCAGCGCGCGCTGCGCTTATCCGGTAACTATGCTTGTAGTCCAAACC 1027
 Db 3156 ACCCCCGCTTTCAGCGCGCGCTGCGCTTATCCGGTAACTATGCTTGTAGTCCAAACC 3215
 QY 1028 GCTAAGACAGCACTTATCGCCACTGCGCAGCAGCCACTGGTAAACAGGATTAGCAGCGAG 1087
 Db 3216 GCTAAGACAGCACTTATCGCCACTGCGCAGCAGCCACTGGTAAACAGGATTAGCAGCGAG 3275
 QY 1088 GTATGTAGCGGTGCTACAGAGTTCCTTAAAGTGGTGGCTTAACTACGCTACACTAGAAG 1147
 Db 3276 GTATGTAGCGGTGCTACAGAGTTCCTTAAAGTGGTGGCTTAACTACGCTACACTAGAAG 3335
 QY 1148 GACAGTATTTGGTATCTGCGCTCTGCTGAAGCCAGTTACCTTCGGAAGAGAGTTGGTAG 1207
 Db 3336 GACAGTATTTGGTATCTGCGCTCTGCTGAAGCCAGTTACCTTCGGAAGAGAGTTGGTAG 3395
 QY 1208 CTCTTGATCCGCAAAACAAACCCGCTGCTGAGCGGTGGTGTGTTTGTTCGAAGCAGCA 1267
 Db 3396 CTCTTGATCCGCAAAACAAACCCGCTGCTGAGCGGTGGTGTGTTTGTTCGAAGCAGCA 3455
 QY 1268 GATTACGCGCAGAAAAAAGGATCT 1292
 Db 3456 GATTACGCGCAGAAAAAAGGATCT 3480
 RESULT 9
 US-08-480-120-20
 ; Sequence 20, Application US/08480120
 ; GENERAL INFORMATION:
 ; APPLICANT: Joliffe, Linda K.
 ; APPLICANT: Zivin, Robert A.
 ; APPLICANT: Pulito, Virginia L.
 ; TITLE OF INVENTION: CDR-GRAFTED ANTI-TISSUE FACTOR
 ; TITLE OF INVENTION: ANTIBODIES AND METHODS OF USE THEREOF
 ; NUMBER OF SEQUENCES: 20
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: Scully, Scott, Murphy & Presser
 ; STREET: 400 Garden City Plaza
 ; CITY: Garden City
 ; STATE: New York
 ; COUNTRY: United States
 ; ZIP: 11530
 ; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: Floppy disk

Query Match	33.4%	Score 604.2	DB 8	Length 7864	
Best Local Similarity	97.9%	Pred. No. 1.8e-167			
Matches 612	Conservative	0	Mismatches 13	Indels 0	Gaps 0
QY	668	TGAGAAAGAAATCTCGACGCGCGCGCTGCTGCGCTTTTTCATAGGCTCGCGCCCC	727		
DB	1255	TAAGGGAGAGCGTCGACCTCGGCGCGCTGCTGCGCTTTTTCATAGGCTCGCGCCCC	1314		
QY	728	TGAGGAGCATCAGAAAATCGACGCTCAAGTCAGAGGTGGCGAAACCCGACAGACTATA	787		
DB	1315	TGACGAGCATCAGAAAATCGACGCTCAAGTCAGAGGTGGCGAAACCCGACAGACTATA	1374		
QY	788	AAGATACCAAGCGCTTTCCGCCCTGGAAGCTCCCTCGTGGCTCTCCTGTTCCGACCCCTGCC	847		
DB	1375	AAGATACCAAGCGCTTTCCGCCCTGGAAGCTCCCTCGTGGCTCTCCTGTTCCGACCCCTGCC	1434		
QY	848	GCTTACCGGATACCTGTCGCCCTTCTCCCTTCGCGGAAGCGTGGCGTTTCTCAATGCTC	907		
DB	1435	GCTTACCGGATACCTGTCGCCCTTCTCCCTTCGCGGAAGCGTGGCGTTTCTCAATGCTC	1494		
QY	908	ACGCTGTAGGTATCTCAGTTTCGGTGTAGTCTCGCTCCAGCTGGCGTGTGTGCACGA	967		
DB	1495	ACGCTGTAGGTATCTCAGTTTCGGTGTAGTCTCGCTCCAGCTGGCGTGTGTGCACGA	1554		
QY	968	ACCCCGCGTTACGCCGACCGCTGCGCCCTTATCCGGTAACTATCGTCTTGAGTCCCAACC	1027		
DB	1555	ACCCCGCGTTACGCCGACCGCTGCGCCCTTATCCGGTAACTATCGTCTTGAGTCCCAACC	1614		
QY	1028	GGTAAACACGACTTATCGCCACTGGCAGCAGCCACTGGTTAAACAGGATATAGCAGAGCGAG	1087		
DB	1615	GGTAAACACGACTTATCGCCACTGGCAGCAGCCACTGGTTAAACAGGATATAGCAGAGCGAG	1674		
QY	1088	GTATGTAGCGCGTGCTACAGAGTTCCTGAAGTGGTGGCCCTAACTACGCGTACACTAGNAG	1147		
DB	1675	GTATGTAGCGCGTGCTACAGAGTTCCTGAAGTGGTGGCCCTAACTACGCGTACACTAGNAG	1734		
QY	1148	GACAGTATTTGGTATCTGCGCTCTGCTGAAGCCAGTTACCTTCGAAAAAGAGTTGGTAG	1207		
DB	1735	GACAGTATTTGGTATCTGCGCTCTGCTGAAGCCAGTTACCTTCGAAAAAGAGTTGGTAG	1794		
QY	1208	CTCTTGATCCGGCAAAACAAACACCGCTGGTGTAGCGGTGTTTTTTTGTGTGCAACGACGA	1267		
DB	1795	CTCTTGATCCGGCAAAACAAACACCGCTGGTGTAGCGGTGTTTTTTTGTGTGCAACGACGA	1854		
QY	1268	GATTACCGCGCAGAAAAAAGGATCT	1292		

	Query Match	33.4%;	Score 603;	DB 16;	Length 608;	
	Best Local Similarity	100.0%;	Pred. No. 1.3e-167;			
	Matches 603;	Conservative	0;	Mismatches	0;	Gaps
QY	690	GC	CGCGCTTGCTGCGCTTTTTCATAGCGTCCGCGCCCGCTTGAGGAGCATCACAAAAATCGA	749		
Db	6	GC	CGCGTGTGCTGCGCTTTTTCATAGCGTCCGCGCCCGCTTGAGGAGCATCACAAAAATCGA	65		
QY	750	CG	CTCAAGTCAGAGTGGCGAACCACGACAGGACTATAAGATACCAAGCGTTTCCCGCCT	809		
Db	66	CG	CTCAAGTCAGAGTGGCGAACCACGACAGGACTATTAAGATACCAAGCGTTTCCCGCCT	125		
QY	810	GGA	AGCTCCCTCGTGGCGCTCTCTGTTCCGACCGCTGCCGTTTACCGGATACCTGTCGCGC	869		
Db	126	GGA	AGCTCCCTCGTGGCGCTCTCTGTTCCGACCGCTGCCGTTTACCGGATACCTGTCGCGC	185		
QY	870	TTT	CTCCCTTCGGGAACGCGTGGCGCTTTCATAGCTCACGCTGTAGGTATCTCAGTTCG	929		
Db	186	TTT	CTCCCTTCGGGAACGCGTGGCGCTTTCATAGCTCACGCTGTAGGTATCTCAGTTCG	245		
QY	930	GTG	TGAGTCTGCTCGCTCCAAAGCTGGGCTGTGTCACGAACCCCCCGTTACGCCGACCGC	989		

Db 246 GTGAGTCTGTTCCAGCTGGCTGTGTGCAGAACCCCGCTTCAGCCGACCCG 305
QY 990 TCGGCTTATCCGGTAACATATCTGTAGTCCAAACCCGGTAAGACACACTATTCGCCA 1049
Db 306 TCGGCTTATCCGGTAACATATCTGTAGTCCAAACCCGGTAAGACACACTATTCGCCA 365
QY 1050 CTGGCAGCAGCCACTGGTAACAGGATTAGCAGAGGAGGTATGTAGCGGTGCTACAGAG 1109
Db 366 CTGGCAGCAGCCACTGGTAACAGGATTAGCAGAGGAGGTATGTAGCGGTGCTACAGAG 425
QY 1110 TTCTTGAAGTGTGGCTTAACATACGGCTACACTAGAGGACAGTATTTGGTATCTGCGCT 1169
Db 426 TTCTTGAAGTGTGGCTTAACATACGGCTACACTAGAGGACAGTATTTGGTATCTGCGCT 485
QY 1170 CTGCTGAAGCCAGTTACCTTCGAAAAAGAGTTGGTAGCTCTTGATCCGGCAACAAACC 1229
Db 486 CTGCTGAAGCCAGTTACCTTCGAAAAAGAGTTGGTAGCTCTTGATCCGGCAACAAACC 545
QY 1230 ACCGCTGTAGCGGTGGTTTTTTTGTTCGAAGCAGCAGATACCGCGCAAAAAAGGA 1289
Db 546 ACCGCTGTAGCGGTGGTTTTTTTGTTCGAAGCAGCAGATACCGCGCAAAAAAGGA 605
QY 1290 TCT 1292
Db 606 TCT 608

RESULT 11

US-09-242-202a-26
; Sequence 26, Application US/09242202a
; GENERAL INFORMATION:
; APPLICANT: Nelson, Edward L.
; TITLE OF INVENTION: NOVEL VECTOR FOR
; POLYNUCLEOTIDE VACCINES
; NUMBER OF SEQUENCES: 29
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: MORGAN & FINNEGAN, L.L.P.
; STREET: 345 PARK AVENUE
; CITY: NEW YORK
; STATE: NEW YORK
; COUNTRY: USA
; ZIP: 10154
; COMPUTER READABLE FORM:
; MEDIUM TYPE: DISKETTE, 3.5 INCH, 1.44 MB
; COMPUTER: IBM PC COMPATIBLE
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: MS WORD 97
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/242, 202a
; FILING DATE: 20-Apr-2000
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US60/023931
; FILING DATE: 14-AUG-1996
; ATTORNEY/AGENT INFORMATION:
; NAME: KATHRYN M. BROWN
; REGISTRATION NUMBER: 34556
; REFERENCE/DOCKET NUMBER: 2026-42360S1
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 758-4800
; TELEFAX: (212) 751-6849
; TELEX: 421792
; INFORMATION FOR SEQ ID NO: 26:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 608 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: unknown
; MOLECULE TYPE: cDNA
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
; SEQUENCE DESCRIPTION: SEQ ID NO: 26:

US-09-242-202a-26

Query Match 33.4%; Score 603; DB 16; Length 608;
Best Local Similarity 100.0%; Pred. No. 1.3e-167; Indels 0; Gaps 0;
Matches 603; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 690 GCGCGCTTCTGCGGCTTTTTCATAGGCTCCGCCCTCCAGAGCATCACAAAAATCGA 749
Db 6 GCGCGCTTCTGCGGCTTTTTCATAGGCTCCGCCCTCCAGAGCATCACAAAAATCGA 65
QY 750 CGCTCAAGTCAGAGGTGGGAAACCCGACAGGACTATAAAGATACACAGCGTTTCCCCCT 809
Db 66 CGCTCAAGTCAGAGGTGGGAAACCCGACAGGACTATAAAGATACACAGCGTTTCCCCCT 125
QY 810 GGAAGCTCCCTGTCGGCTCTCTCTGTTCGGACCTCCCGGTACCCGATACCTGTCGGCC 869
Db 126 GGAAGCTCCCTGTCGGCTCTCTCTGTTCGGACCTCCCGGTACCCGATACCTGTCGGCC 185
QY 870 TTTCTCCCTTCGGAAGCGTGGCGCTTTCTCAATGCTCACGCTGTAGGTATCTCAGTTCTG 929
Db 186 TTTCTCCCTTCGGAAGCGTGGCGCTTTCTCAATGCTCACGCTGTAGGTATCTCAGTTCTG 245
QY 930 GTGTAGTCTGCTTCCAAAGCTGGGCTGTGTGCACGAACCCCGCTTCAGCCCGACCCG 989
Db 246 GTGTAGTCTGCTTCCAAAGCTGGGCTGTGTGCACGAACCCCGCTTCAGCCCGACCCG 305
QY 990 TCGGCTTATCCGGTAACATATCTGTAGTCCAAACCCGGTAAGACACACTATTCGCCA 1049
Db 306 TCGGCTTATCCGGTAACATATCTGTAGTCCAAACCCGGTAAGACACACTATTCGCCA 365
QY 1050 CTGGCAGCAGCCACTGGTAACAGGATTAGCAGAGGAGGTATGTAGCGGTGCTACAGAG 1109
Db 366 CTGGCAGCAGCCACTGGTAACAGGATTAGCAGAGGAGGTATGTAGCGGTGCTACAGAG 425
QY 1110 TTCTTGAAGTGTGGCTTAACATACGGCTACACTAGAGGACAGTATTTGGTATCTGCGCT 1169
Db 426 TTCTTGAAGTGTGGCTTAACATACGGCTACACTAGAGGACAGTATTTGGTATCTGCGCT 485
QY 1170 CTGCTGAAGCCAGTTACCTTCGAAAAAGAGTTGGTAGCTCTTGATCCGGCAACAAACC 1229
Db 486 CTGCTGAAGCCAGTTACCTTCGAAAAAGAGTTGGTAGCTCTTGATCCGGCAACAAACC 545
QY 1230 ACCGCTGTAGCGGTGGTTTTTTTGTTCGAAGCAGCAGATACCGCGCAAAAAAGGA 1289
Db 546 ACCGCTGTAGCGGTGGTTTTTTTGTTCGAAGCAGCAGATACCGCGCAAAAAAGGA 605
QY 1290 TCT 1292
Db 606 TCT 608

RESULT 12

PCT-US99-06742-7
; Sequence 7, Application PC/TUS9906742
; GENERAL INFORMATION:
; APPLICANT: Ruffner, Duane E.
; APPLICANT: Pierce, Michael L.
; APPLICANT: Chen, Zhidong
; TITLE OF INVENTION: Directed Antisense Libraries
; FILE REFERENCE: T6678.PCT
; CURRENT APPLICATION NUMBER: PCT/US99/06742
; CURRENT FILING DATE: 1999-03-28
; EARLIER APPLICATION NUMBER: 60/079,792
; EARLIER FILING DATE: 1998-03-28
; NUMBER OF SEQ ID NOS: 34
; SOFTWARE: Word97
; SEQ ID NO: 7
; LENGTH: 2077
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: paslib

PCT-US99-06742-7

Query Match 33.4%; Score 603; DB 1; Length 2077;
Best Local Similarity 100.0%; Pred. No. 2.3e-167;
Matches 603; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 690 GCCGCGTGTGCGCTTTTTCATAGGCTCCGCCCTTGACGAGCATCACAAAATCGA 749
DB 1437 gcgcgtgtgctgttcttcataagctccgcccttgacgagcatcaaaaaatcga 1496
QY 750 CCCTCAAGTCAGAGTGGCGAAACCCGACAGACTATAAAGATACAGCGCTTTCCGCCCT 809
DB 1497 cgcctcaagtcagagtgccgaaacccgacagactataaagataccagcgcttcccccct 1556
QY 810 GGAAGCTCCCTCGCTGCTCTCTCTGTCGACCGCTTCAATGCTACGCTGTAGTATCTCAGTTTCG 929
DB 1557 ggaagctccctcgctgctctctctgttcgacacctgcccgtttaccggatacctgtccgcc 1616
QY 870 TTTCTCCCTTCGGGAAGCTGGCGCTTTCTCAATGCTACGCTGTAGTATCTCAGTTTCG 929
DB 1617 ttctcccttcgggaagcgtggcgcttctcctcaatgctcaagctgtaggtatctcagttcg 1676
QY 930 GTGTAGTGTCTTCGCTTCCAAAGCTGGGCTGTGTGACGACGACCCCGCTTCAGCCGACCGC 989
DB 1677 gtgtaggtgtcttcgctccaaagctggctgtgtgacgaaccccccgcttcagccgacgcg 1736
QY 990 TGGCGCTTATCCGGTAACTATCGCTTGTAGTCCAAACCCGGTAAGACAGCACTTATCGCCA 1049
DB 1737 tggccttatccggtaactatgctgttgatccaaacccggtaagacagacttatcgcca 1796
QY 1050 CTGGCAGCAGCCACTGTAACAGGATTAGCAGAGCGAGGATTAGTGGCGGTGCTACAGAG 1109
DB 1797 ctggcagcagccactggttaacaggaattagcagagcgaggtatgtaggcggtgctacagag 1856
QY 1110 TTTCTGAAGTGTGGGCTAACTACGCTACACTAGAGACAGACTATTGGTATCTCGGCT 1169
DB 1857 ttctgaagtggtggcctaactacgctacactagagagacagattttggtatctcgct 1916
QY 1170 CTGCTGAAGCCAGTTACCTTCGAAAAGAGTTGGTAGCTCTTGATCCGGCAAAACACC 1229
DB 1917 ctgctgaagccagttaccttcgaaaagagttggtagctcttgatccggcaaaacacc 1976
QY 1230 ACCGCTGTAGCGGTGGTGTGTTTGGTTCGACGAGCATTTACGCGCAAAAAGGA 1289
DB 1977 accgctggtagcggtgtgtttttgttgcagcagcagatttacgcgcgcaaaaaaagga 2036
QY 1290 TCT 1292
DB 2037 tct 2039

RESULT 13
US-09-496-445-5
; Sequence 5, Application US/09496445
; GENERAL INFORMATION:
; APPLICANT: McNeish, John D.
; APPLICANT: Ahljinian, Michael K
; TITLE OF INVENTION: Transgenic Animals Expressing Human p25
; FILE REFERENCE: PC10142A
; CURRENT APPLICATION NUMBER: US/09/496,445
; CURRENT FILING DATE: 2000-02-02
; NUMBER OF SEQ ID NOS: 7
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 5
; LENGTH: 2462
; TYPE: DNA
; ORGANISM: Escherichia coli
US-09-496-445-5

Query Match 33.4%; Score 603; DB 18; Length 2462;
Best Local Similarity 100.0%; Pred. No. 2.5e-167;

Matches 603; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 690 GCCGCGTGTGCTGGCTTTTTCATAGGCTCCGCCCTTGACGAGCATCACAAAATCGA 749
DB 359 gcgcgtgtgctgttcttcataagctccgcccttgacgagcatcaaaaaatcga 418
QY 750 CGCTCAAGTCAGAGTGGCGAAACCCGACAGACTATAAAGATACAGCGGTTCGCCCT 809
DB 419 cgcctcaagtcagagtgccgaaacccgacagactataaagataccagcgcttcccccct 478
QY 810 GGAAGCTCCCTCGCTGCTCTCTGTCGACCGCTTCAATGCTACGCTGTAGTATCTCAGTTTCG 869
DB 479 ggaagctccctcgctgctctctctgttcgacacctgcccgtttaccggatacctgtccgcc 538
QY 870 TTTCTCCCTTCGGGAAGCTGGCGCTTTCTCAATGCTACGCTGTAGTATCTCAGTTTCG 929
DB 539 ttctcccttcgggaagcgtggcgcttctcctcaatgctcaagctgtaggtatctcagttcg 598
QY 930 GTGTAGTGTCTTCGCTTCCAAAGCTGGGCTGTGTGACGACGACCCCGCTTCAGCCGACCGC 989
DB 599 gtgtaggtgtctcctcaagctggcgctgtgtgcaagaccccccgcttcagccgacgcg 658
QY 990 TGGCGCTTATCCGGTAACTATCGCTTGTAGTCCAAACCCGGTAAGACAGCACTTATCGCCA 1049
DB 659 tggccttatccggtaactatgctgttgatccaaacccggtaagacagacttatcgcca 718
QY 1050 CTGGCAGCAGCCACTGTAACAGGATTAGCAGAGCGAGGATTAGTGGCGGTGCTACAGAG 1109
DB 719 ctggcagcagccactggttaacaggaattagcagagcaggtatgtaggcggtgctacagag 778
QY 1110 TTTCTGAAGTGTGGGCTAACTACGCTACACTAGAGACAGACTATTGGTATCTCGGCT 1169
DB 779 ttctgaagtggtggcctaactacgctacactagagagacagattttggtatctcgct 838
QY 1170 CTGCTGAAGCCAGTTACCTTCGAAAAGAGTTGGTAGCTCTTGATCCGGCAAAACACC 1229
DB 839 ctgctgaagccagttaccttcgaaaagagttggtagctcttgatccggcaaaacacc 898
QY 1230 ACCGCTGTAGCGGTGGTGTGTTTGGTTCGACGAGCATTTACGCGCAAAAAGGA 1289
DB 899 accgctggtagcggtgtgtttttgttgcagcagcagatttacgcgcgcaaaaaaagga 958
QY 1290 TCT 1292
DB 959 tct 961

Query Match 33.4%; Score 603; DB 44; Length 2462;
Best Local Similarity 100.0%; Pred. No. 2.5e-167;
Matches 603; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 690 GCCGCGTGTGCGGTGTTTTCATAGGCTCCGCCCTTGACGAGCATCACAAAATCGA 749
DB 359 gcgcgtgtgctgttcttcataagctccgcccttgacgagcatcaaaaaatcga 418

|||||
Db 1276 ACCGCTGCTAGCGTGGTTTTTTTGTTCAGCAGCAGATTACCGCAGAAAAAGGA 1335
QY 1290 TCT 1292
|||
Db 1336 TCT 1338

Search completed: January 17, 2002, 15:14:36
Job time: 22788 sec